

**УНИВЕРЗИТЕТ „СВ. КЛИМЕНТ ОХРИДСКИ“**



**ФАКУЛТЕТ ЗА ИНФОРМАТИЧКИ И КОМУНИКАЦИСКИ  
ТЕХНОЛОГИИ - БИТОЛА**

**- КАТЕДРА ЗА ТЕОРЕТСКИ ОСНОВИ НА ИНФОРМАТИКАТА И ПРЕСМЕТКОВНО  
ИНЖЕНЕРСТВО -**

**ДОКТОРСКИ ТРУД**

НА ТЕМА

**Бихејвиористички-базирано моделирање на  
процесите на пребарување и навигација низ  
информациски мрежи**

Ментор

Проф. Д-р Пеце Митревски

Кандидат

М-р Весна Кумбароска

Битола, 2015

**ЧЛЕНОВИ НА КОМИСИЈАТА ЗА ОЦЕНКА И ОДБРАНА НА ДОКТОРСКАТА ДИСЕРТАЦИЈА:**

**Д-Р ПЕЦЕ МИТРЕВСКИ (Ментор)**

РЕДОВЕН ПРОФЕСОР

ФАКУЛТЕТ ЗА ИНФОРМАТИЧКИ И КОМУНИКАЦИСКИ ТЕХНОЛОГИИ

УНИВЕРЗИТЕТ „СВ. КЛИМЕНТ ОХРИДСКИ“ - БИТОЛА

**Д-Р ВИОЛЕТА МАНЕВСКА (Претседател)**

РЕДОВЕН ПРОФЕСОР

ФАКУЛТЕТ ЗА ИНФОРМАТИЧКИ И КОМУНИКАЦИСКИ ТЕХНОЛОГИИ

УНИВЕРЗИТЕТ „СВ. КЛИМЕНТ ОХРИДСКИ“ - БИТОЛА

**Д-Р ИЛИЈА ЈОЛЕВСКИ (Член)**

ВОНРЕДЕН ПРОФЕСОР

ФАКУЛТЕТ ЗА ИНФОРМАТИЧКИ И КОМУНИКАЦИСКИ ТЕХНОЛОГИИ

УНИВЕРЗИТЕТ „СВ. КЛИМЕНТ ОХРИДСКИ“ - БИТОЛА

**Д-Р АЛЕКСАНДАР МАРКОВСКИ (Член)**

РЕДОВЕН ПРОФЕСОР

ФАКУЛТЕТ ЗА ИНФОРМАТИЧКИ И КОМУНИКАЦИСКИ ТЕХНОЛОГИИ

УНИВЕРЗИТЕТ „СВ. КЛИМЕНТ ОХРИДСКИ“ - БИТОЛА

**Д-Р ЛЪПЧО АНТОВСКИ (Член)**

ВОНРЕДЕН ПРОФЕСОР

ФАКУЛТЕТ ЗА ИНФОРМАТИЧКИ НАУКИ И КОМПЈУТЕРСКО ИНЖЕНЕРСТВО

УНИВЕРЗИТЕТ „СВ. КИРИЛ И МЕТОДИЈ“ - СКОПЈЕ

## Декларација

---

Потврдувам дека, освен за случаи каде е направено прописно цитирање, оваа докторска теза со наслов “Бихејвиористички-базирано моделирање на процесите на пребарување и навигација низ информациски мрежи” е оригинално дело на самиот автор; тезата не е поднесена претходно, во целост или делумно, да се квалификува за каква било друга академска награда; содржината на тезата е резултат на работата која е извршена од официјалниот датум на одобрување на пријавата; и, секоја уредувачка работа, платена или неплатена, извршена од страна на трети лица е прописно референцирана.

Докторската теза е напишана во (Word 2013) на (Windows 7) платформа.

---

Весна Кумбароска

Факултет за информатички и комуникациски технологии

Катедра за теоретски основи на информатиката и пресметковно инженерство

Битола, Септември 2015

## Благодарност

---

Сакам особено да се заблагодарам на мојата фамилија за непоколебливата поддршка и разбирање за многуте моменти кои јас ги посветив на постигнување на оваа пресвртница во мојот живот и кариера.

Исто така, исклучителна благодарност до мојот менторот проф. Д-р Пеце Митревски кој несебично ми пружи професионална соработка и своето знаење целосно го споделуваше во заедничката работа.

На крај, сакам да изразам голема благодарност до тимот на електронската книжарница [www.kupikniga.mk](http://www.kupikniga.mk) кој излезе во пресрет на оваа истражувачка работа и несебично сподели дел од своите приватни податоци, без што не можеше навремено и квалитетно да се реализира оваа докторска дисертација..

## Апстракт

---

Посетата и пазарувањето од електронските продавници е сè почеста активност на денешните интернет корисници. Сведоци сме на секојдневно еволуирање на овие сервиси во насока на прилагодување кон корисниците т.е. нивното однесување, преку: предвидување и предлагање на слични содржини или производи кои претходно биле погледнати или купени, персонализација, таргетиран маркетинг итн., со цел, од една страна, зголемување на продажбата и претворање на што поголем дел пребарувачи во купувачи, забрзување и зголемување на можноста на избор од што поголемо множество продукти, задржување на вниманието на купувачите, нивната лојалност и сл., а од друга страна, прилагодување и подобрување на веб дизајнот, односно интерфејсот, со цел подобрување на корисничкото искуство и интеракцијата со системот.

Во целата оваа слика, освен анализата на корисничкото однесување при процесот на купување, разбирањето на целокупното корисничко однесување, е важен показател на корисничките интереси. Во таа насока, оваа докторска дисертација, ги идентификуваше најчестите кориснички шеми на однесување при процесот на пребарување и навигација, како нераздвојни делови од процесот на електронско купување. Предложени се два бихејвиористички-базирани модели кои даваат продлабочени знаења за процесите на пребарување и навигација низ информациски мрежи илустрирано низ примерот на електронска книжарница (продавница). Развиените модели се засновани на стохастичко моделирање со помош на Петриевите мрежи, во чија основа на нивно решавање лежи Марков Ланец во Континуирано Време, со чија помош ги евалуиравме моделите и пресметавме одредени мерки на перформанси, како што се: очекувано време поминато во преодно допирливо означување, кумулативно време на престој во преодно допирливо означување, вкупен број на посети во преодно допирливо означување и сл. Исто така, илустриравме директна пресметка на некои компоненти на ангажираност на корисникот и предложивме употреба на развиениот бихејвиористички-базиран модел на навигација за мерење на ангажираност на корисникот.

## Abstract

---

Nowadays, online shopping is becoming very frequent user activity and we are witnesses of daily evolution of these kind of services in several directions adjusting to the user behavior, such as: predicting and proposing similar content or products that were previously previewed or purchased, personalization, targeted marketing, etc. On one hand, this leads to increased sales and turning the most browsers into buyers, speeded up and increased possibility of choosing of a big set of products, kept costumers attention, loyalty, etc., and on the other hand, adjustment and improvement of the web design or the interface in order to improve the user experience and interaction with the system.

In this whole picture, besides the analysis of the user behavior in the electronic purchasing process, understanding the overall user behavior is an important indicator of user's interests. In this regard, this doctoral dissertation has identified the most common user behavior patterns in the processes of searching and navigation, as inseparable parts of the electronic purchasing proces. Two behavioral-based models that provide profound knowledge about the processes of searching and navigation through information networks illustrated through the example of an electronic bookstore are proposed. The developed models are based on stochastic modeling using Petri Nets, which solution is based on Continious Time Markov Chain, used for model evaluation and few performance masures calculation, such as: the expected time spent in a transient tangible marking, cumulative sojourn time spent in a transient tangible marking, total number of visits in a transient tangible marking etc. Also, direct calculation of some user engagement components is illustrated and usage of the developed navigational behavioral-based model to measure user engagement is proposed.

## Кратка содржина

---

<b>1</b>	<b>ВОВЕД</b> .....	<b>2</b>
1.1	ПРЕДМЕТ И ОБЛАСТ НА ИСТРАЖУВАЊЕ.....	3
1.2	ЦЕЛИ НА ИСТРАЖУВАЊЕТО .....	4
1.3	ИСТРАЖУВАЧКИ ПРАШАЊА .....	5
1.4	МЕТОДОЛОГИЈА НА ИСТРАЖУВАЊЕ.....	6
1.5	СТРУКТУРА НА ДИСЕРТАЦИЈАТА .....	6
<b>2</b>	<b>ТЕОРИСКИ ОСНОВИ</b> .....	<b>9</b>
2.1	ГЕНЕРАЛИЗИРАНИ СТОХАСТИЧКИ ПЕТРИЕВИ МРЕЖИ .....	11
2.2	МАРКОВ ЛАНЕЦ ВО КОНТИНУИРАНО ВРЕМЕ .....	12
2.3	КОРИСТЕЊЕ НА GSPN МРЕЖИ ЗА МОДЕЛИРАЊЕ И ЕВАЛУАЦИЈА НА ПЕРФОРМАНСИ .....	13
2.4	ЗАВРШЕН ПРЕГЛЕД НА ПОГЛАВЈЕТО .....	16
<b>3</b>	<b>АКТУЕЛНОСТ НА ИСТРАЖУВАЊЕТО</b> .....	<b>18</b>
3.1	МОДЕЛИРАЊЕ НА ОДНЕСУВАЊЕТО НА КОРИСНИЦИТЕ ПРИ ПРЕБАРУВАЊЕ .....	18
3.2	МОДЕЛИРАЊЕ НА РЕФОРМУЛИРАЊЕТО НА ПРАШАЊЕТО ПРИ ПРЕБАРУВАЊЕ .....	19
3.3	МОДЕЛИРАЊЕ НА ОДНЕСУВАЊЕТО НА КОРИСНИЦИТЕ ПРИ НАВИГАЦИЈА.....	19
3.4	МЕРЕЊЕ НА НЕМЕРЛИВОТО: АНГАЖИРАНОСТ НА КОРИСНИКОТ .....	21
3.5	КЛАСТЕРИРАЊЕ НА СЕКВЕНЦИ ОД ПРЕМИНИ.....	22
3.6	ЗАВРШЕН ПРЕГЛЕД НА ПОГЛАВЈЕТО .....	23
<b>4</b>	<b>ПРЕДЛОЖЕНИ МОДЕЛИ</b> .....	<b>25</b>
4.1	КОРИСТЕЊЕ НА GSPN МРЕЖИ ЗА ОТСЛИКУВАЊЕ НА ШЕМИТЕ НА ОДНЕСУВАЊЕ ПРИ ПРЕБАРУВАЊЕ.....	25
4.1.1	<i>Анализа на предложениот проширен модел</i> .....	30
4.1.2	<i>Анализа на стабилни состојби и мерење на перформанси</i> .....	31
4.1.3	<i>Евалуација на мерки на перформанси преку студија на случај</i> .....	32
4.2	КОРИСТЕЊЕ НА GSPN МРЕЖИ ЗА ОТСЛИКУВАЊЕ НА ШЕМИТЕ НА ОДНЕСУВАЊЕ ПРИ НАВИГАЦИЈА .....	35
4.2.1	<i>Предложен модел</i> .....	35
4.2.2	<i>Анализа на предложениот модел</i> .....	47
4.2.3	<i>Анализа на стабилни состојби и мерење на перформанси</i> .....	52
4.2.4	<i>Евалуација на мерки на перформанси преку студија на случај</i> .....	54
4.2.4.1	Кластерирање на секвенци – концепт и примена .....	58
4.2.4.2	Интерпретирање на кластерите .....	60

4.2.4.3	Решавање на моделот .....	67
4.2.5	Завршен преглед на поглавјето .....	74
4.2.6	Идни правци .....	75
<b>5</b>	<b>ПРИЛОГ КОН ПРЕСМЕТКА НА КОМПОНЕНТИТЕ НА АНГАЖИРАНОСТ НА КОРИСНИКОТ .....</b>	<b>77</b>
5.1	КОРИСТЕЊЕ НА ЛОГ ПОДАТОЦИТЕ ЗА ПРЕСМЕТКА НА КОМПОНЕНТИТЕ НА АНГАЖИРАНОСТ НА КОРИСНИКОТ.....	78
5.2	КОРИСТЕЊЕ НА БИХЕЈВИОРИСТИЧКИ-БАЗИРАНИТЕ МОДЕЛИ ЗА ПРЕСМЕТКА НА КОМПОНЕНТИТЕ НА АНГАЖИРАНОСТ НА КОРИСНИКОТ .....	89
5.3	ЗАВРШЕН ПРЕГЛЕД НА ПОГЛАВЈЕТО .....	90
<b>6</b>	<b>ЗАКЛУЧОК .....</b>	<b>92</b>
	<b>БИБЛИОГРАФИЈА.....</b>	<b>94</b>
	<b>ЛИСТА СО ТАБЕЛИ.....</b>	<b>101</b>
	<b>ЛИСТА СО СЛИКИ .....</b>	<b>103</b>
	<b>ДОДАТОК А – РЕСУРСИ.....</b>	<b>105</b>
	<b>ДОДАТОК Б – РАБОТНА ОКОЛИНА .....</b>	<b>106</b>
	<b>ДОДАТОК Ц – СТАТИСТИЧКИ АНАЛИЗИ И ЕВАЛУАЦИЈА НА ПЕРФОРМАНСИ .....</b>	<b>107</b>
	<b>КРАТЕНКИ И АКРОНИМИ .....</b>	<b>111</b>



## ***Глава 1***

## 1 Вовед

---

Во поново време сè повеќе е зголемена популарноста на електронските продавници од различен тип, што води кон зголемен број на корисници (купувачи) и пофреквентно онлајн купување. Исто така се поставуваат нови стандарди и се зголемуваат очекувањата на корисниците. Тоа подразбира потреба за враќање на што поточни резултати при процесот на пребарување (information retrieval), поточни препораки (recommendation), подобар дизајн ориентиран кон корисникот (user experience design), како и потреба за што подобра персонализација (personalization), како при процесот на пребарување, така и при процесот на навигација, во различни сфери, меѓу кои и сферата на електронската трговија. Сето ова бара продлабочено разбирање на интеракцијата, односно, на интерфејсот од една страна, и однесувањето на корисниците во таквото електронско опкружување, од друга страна, што е исклучително важно за креирање на нови стратегии и соодветен маркетинг.

Сè побројни се истражувањата кои се фокусираат на откривање и анализа на сложениот профил на корисникот (купувачот) и неговото однесување. Стратувајќи од оваа точка и гледајќи ја оваа област интересна и предизвикувачка, реализиравме истражување и испитување дали Петриевите мрежи, како модели базирани врз транзиции (премини), можат да бидат применети за **анализа и моделирање** на однесувањето на корисниците при пребарување и навигација низ информациски мрежи.

Во таа насока е развиена оваа докторска дисертација, со која се дава директен придонес кон пополнувањето на една битна празнина во оваа област. Ќе се спроведе темелно и опсежно научно истражување во правец на развој на стохастички модели за однесување на корисниците при пребарување и навигација низ информациски мрежи. Под претпоставка дека корисничкото однесување може да биде опишано како стохастички процес со експоненцијално дистрибуирани временски интервали, а корисничките сесии (посети) имаат експоненцијално дистрибуирано траење со средна вредност еднаква на кумулативното време на престој, преку развиените модели лесно ќе можеме да ја опишеме “оперативната околина”, т.е. да го поделиме клиентскиот “влезен простор” преку групирање на корисниците кои имаат што поблиски шеми на однесување, како и да се направи евалуација на доверливоста на сервисот, достапноста и перформансите на системот, и посебен акцент да биде ставен на ангажираноста на корисниците. Разбирањето на корисничките акции може да помогне во подобро прилагодување на сè што претставува хаос, поточно на располагање да се стават оние можности кои го поддржуваат спектарот на веќе познати шеми на однесување. Вака дефинираната област,

не е доволно истражувана во светот, а со оглед на фактот што кај нас е во зачеток, природно се наметнува како актуелна и погодна за реализација на научни истражувања.

## 1.1 Предмет и област на истражување

Посетата и пазарувањето од електронските продавници е сè почеста активност на денешните интернет корисници. Сведоци сме на секојдневно еволуирање на овие сервиси во насока на: прилагодување кон корисниците т.е. нивното однесување, предвидување и предлагање на слични содржини (производи) кои претходно биле погледнати или купени, таргетиран маркетинг и сл.

Фокусот на нашето истражување е поставен врз откривање и моделирање на однесувањето на корисниците, односно купувачите на книги при процесот на пребарување и навигација, специфично испитано врз првиот случај на електронска книжарница<sup>1</sup> во Република Македонија, но лесно генерализиран и применлив врз познатите сервиси како: : Amazon<sup>2</sup>, Barnes&Noble<sup>3</sup>, Pawell's Books<sup>4</sup> и сл., чија структура на системот е во голема сличност со избраното сценарио.

Целокупниот процес на истражување се базира на следните активности:

1. Пред-процесирање на собраните навигациски податоци од електронската продавница во одредена случајна временска рамка т.е. целокупната интеракција на корисниците со веб сајтот. Тоа подразира чистење на податоците, нивна трансформација и стандардизација со цел полесна и побрза понатамошна обработка и анализа.
2. Статистича анализа на гореспоменатите податоци. Дел од анализите се: што е најинтересно за корисниците, успешно завршени електронски купувања, утврдување на заинтересираноста на корисниците за пребарување по различни критериуми на расположливи производи или пак директен пристап кон конкретен производ, утврдување на корисниците кои се одлучиле да прочитаат примерок на дадена книга и сл.
3. Покрај автоматско утврдување на секој преглед на даден сегмент (веб страна) од веб сајтот, вклучено е и развивање на алгоритам за утврдување на времињата помеѓу две последователни посети, колкаво е во просек тоа време, најдолго односно најкратко време, и сл.
4. Врз основа на животниот циклус на корисничките посети т.е. врз основа на описот на секвенците на премини, одредување на типовите корисници кои се карактеристични за избраната електронска продавница.

---

<sup>1</sup> <http://www.kupikniga.mk>

<sup>2</sup> <http://www.amazon.com/books-used-books-textbooks/b?node=283155>

<sup>3</sup> <http://www.barnesandnoble.com/>

<sup>4</sup> <http://www.powells.com/>

5. Истражувањето продолжува во правец на креирање на модели со употреба на класите на Петриеви мрежи со стохастичко поведење – Генерализирани стохастички Петриеви мрежи (GSPN), за опишување на специфичното однесување на корисниците при навигација, како и генерални модели кои ќе го опишуваат процесот на пребарување. Динамичката природа на Петриевите мрежи, посебно на GSPN мрежите, укажува дека тие лесно можат да бидат прилагодени и применливи за моделирање на однесувањето на корисниците во процесот на пребарување и навигација низ информациски мрежи.
6. Во основата на решавањето на ваквите модели лежи идејата да се проучи нивната еволуција како Марков Ланец (MC) во Континуирано Време (СТМС). Ова води кон можноста за пресметка и евалуација на одредени мерки на перформанси.
7. На крај, истражувањето завршува со директна пресметка на компонентите на ангажираност на корисникот и за **прв пат предлагање употреба** на развиениот модел за мерење на дел од компонентите на ангажираност на корисникот.

Имајќи го предвид погоре изнесеното, ПРЕДМЕТОТ НА ИСТРАЖУВАЊЕ на оваа докторска дисертација може да се дефинира како: истражување на можностите за примена на Петриевите мрежи со стохастичко поведење како алатка за изградба и анализа на модели на однесување на корисниците при пребарување и навигација низ информациски мрежи, демонстрирано низ примерот на електронска продавница.

## 1.2 Цели на истражувањето

Стохастичната природа на корисничката интеракција со системот претставува доволен мотив ова истражување да се насочи токму кон дефинирање на стохастички модели кои во целост ќе соодветствуваат на реалната слика. Со други зборови, однесувањето на корисниците во рамки на еден ваков систем може да се опише како процес на наизменични премини од една во друга состојба, што укажува дека е сосема природно процесот на пребарување и навигација да биде претставен со помош на Петриеви мрежи. Кога зборуваме за концептот состојба, генерално при процесот на пребарување може да се утврдат следните состојби:

1. Формулирање на ново прашање (query)
2. Преглед на резултати кои се вратени
3. Модифицирање на дадено прашање: додавање, бришење и промена на зборови
4. Враќање на претходно прашање
5. Повеќе резултати за дадено прашање (следна страна со резултати)
6. Избор на даден резултат за понатамошен преглед
7. Крај на сесијата

Додека пак дел од утврдените најчести состојби карактеристични за процесот на навигација во рамки на електронска продавница се следните:

1. Преглед на почетна страна
2. Преглед на производи од избрана категорија
3. Преглед на избран производ
4. Пребарување по различни критериуми и преглед на повратните резултати
5. Преглед на корисничката сметка на логираниот корисник
6. Ставање во кошничка
7. Последен преглед на кошничка
8. Купување и плаќање
9. Најпродавани производи
10. Попусти
11. Вести
12. Помош
13. Нови производи
14. Читање на исечок од книга (специфично за нашиот случај) и сл.

Со оглед на сите овие изложени факти, може да се постави следнава ТЕОРЕТСКАТА ЦЕЛ на докторската дисертација: изградба на нови модели за опишување и анализа на однесувањето на корисниците во рамки на електронска продавница, заснована на стохастичко моделирање со помош на Петриевите мрежи, валидација, решавање на ваквите модели, и, врз основа на тоа, евалуација на соодветни мерки на перформанси, и на крај, мерење на ангажираност на корисникот.

### **1.3 Истражувачки прашања**

Водејќи се кон остварување на целите, конкретно, обработени се следните прашања:

1. Креирање на стохастички модели кои реално го опишуваат генералното однесување на корисниците при пребарување, односно реформулирање на прашањето.
2. Карактеризација на типови на корисници во зависност од нивната секвенца на премини при процесот на навигација.
3. Креирање на стохастички модел кој ќе даде реален и детален приказ на однесувањето на специфичните корисници при навигација односно интеракција со електронската продавница, или, генерализирано, низ информациска мрежа.
4. Анализата на ваквите модели, , врз основа на тоа, евалуацијата на соодветни мерки на перформанси е со цел ефективна работа на сервисот електронска продавница и директна пресметка на некои компоненти на ангажираност на

корисникот како и предлог употреба на развиениот бихејвиористички-базиран модел на навигација за мерење на ангажираност на корисникот.

#### 1.4 Методологија на истражување

Методологијата за одговарање на клучните прашања за оставрување на поставените цели може да биде структурирана на следниот начин:

- аналитички и компаративен метод на истражување,
- индуктивен и дедуктивен метод на заклучување,
- методологија на моделирање на системи со примена на Петриевите мрежи со стохастичко поведење,
- системски пристап преку употребата на серија математички и статистички методологии и постапки.

#### 1.5 Структура на дисертацијата

Докторската дисертацијата се состои од повеќе глави, организирани на следниот начин:

Апстракт

- 1 Вовед
- 2 Теориски основи
- 3 Актуелност на истражувањето
- 4 Предложени модели
- 5 Прилог кон пресметката на компонентите на ангажираност на корисникот
- 6 Заклучок и дискусија

Библиографија

Додаток

**Глава 1** претставува вовед во проблематиката на истражување. Елаборирани се мотивите за отпочнување на едно вакво истражување, предметот и целта на истражувањето, како и очекуваните придобивки од него. Во кратки црти објаснета е и самата структурна организација на трудот.

**Глава 2** содржи теоретски основи на оваа проблематика. Акцентот е ставен на концептот на Петриевите мрежи со стохастичко поведење и како истите можат да бидат применети за опишување и анализа на однесувањето на корисниците.

**Глава 3** дава детален преглед на досегашните истражувања во оваа област, достапни во литературата, со посебен акцент на: Петриевите мрежи со стохастичко поведење,

моделирање на однесувањето на корисниците при пребарување, моделирање на реформулирањето на прашањето при пребарување, моделирање на однесувањето на корисниците при навигација, мерење на ангажираност на корисниците, како и постоечките методи за кластерирање на секвенци од премини.

**Глава 4** во целост е посветена на деталниот опис на предложените модели, засновани на теоријата на моделирање со примена на класите на Петриеви мрежи со стохастичко поведење. Елаборирани се основните причини за воведување на ваквите стохастички модели на однесување на корисниците при пребарување и навигација. Даден е детален опис на структурата на моделите, нивно решавање, валидација и, врз основа на тоа, евалуација на соодветни мерки на перформанси со цел ефективна работа на сервисот.

**Глава 5** претставува прилог кон пресметката на компонентите на ангажираност на корисниците, искористувајќи ги лог податоците и развиениот бихејвиористички-базиран модел на процесот на навигација, модифицирај го моделот на Петерсон, така што наместо оригиналниот пристап каде се користат сесии, илустрирана е примена на време помеѓу две последователни посети.

**Глава 6** презентира и дискутира заклучни согледувања во врска со обработуваната тематика, резимира определени резултати и заклучоци од направените истражувања и назначува насоки за понатамошна работа.

## ***Глава 2***



## 2 Теориски основи

---

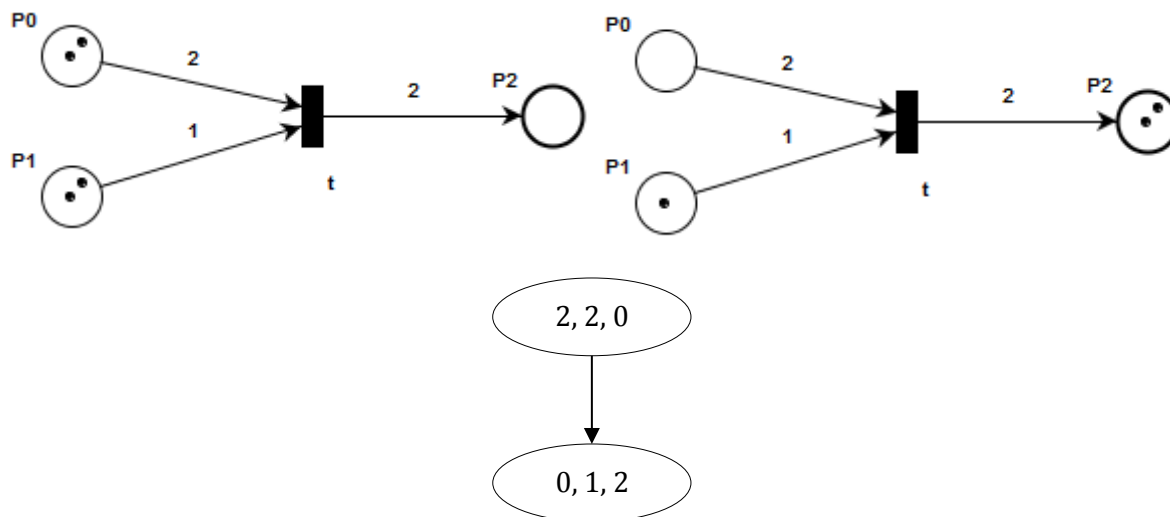
Во ова поглавје ќе биде претставен концептот на Петриевите мрежи чија динамичка природа укажува дека тие можат да бидат прилагодени за бихејвиористички-базирано моделирање на реалното однесување на корисниците при процесите на пребарување и навигација низ информациски мрежи. Ќе биде даден приказ на нивните основни елементи, карактеристики и ќе биде илустрирано од каде потекнува нивната динамика. Посебно внимание ќе биде посветено на Петриевите мрежи со стохастичко поведење, кои се искористени во нашето истражување и во чија основа на решавање лежи Марков Ланец во Континуирано Време.

Формалната дефиниција на Петриевите мрежи [45], [50] укажува дека секоја Петриева мрежа е составена од пет компоненти:  $PN = (P, T, F, W, M_0)$ , така што:

1.  $P = (p_1, p_2, \dots, p_n)$  претставува множество места (places),
2.  $T = (t_1, t_2, \dots, t_n)$  претставува множество премини (transitions),
3.  $F \subseteq (P \times T) \cup (T \times P)$  претставува множество лази (arcs), така што првата компонента се однесува на насочен лак од место кон премин, а втората се однесува на обратниот случај, односно, насочен лак од премин кон место,
4.  $W: F \rightarrow \{1, 2, 3 \dots\}$  претставува тежинска функција,
5.  $M_0: P \rightarrow \{0, 1, 2, 3 \dots\}$  претставува иницијално означување (marking), кое е во релација со терминот жетон (token),
6. Секогаш,  $P \cup T \neq \emptyset$  и  $P \cap T = \emptyset$ .

Во таа насока, може да се каже дека основните елементи на ваквите мрежи се: места, премини, жетони и лази. Најчесто, кога мрежата се претставува графички, местата се претставени како кругови, премините како правоаголници и жетоните како црни точки. Местата се однесуваат на состојби, премините на акции кои може да ги променат состојбите, додека пак лаците ја одредуваат насочената релација помеѓу местата и премините [45], [50]. Означувањата на Петриевите мрежи е тесно поврзано со жетоните и се употребува за опишување на динамичкото однесување на системот. Во тој контекст, преминот е овозможен ако сите влезни места на тој премин содржат најмалку еден жетон. Веќе овозможениот премин може да се активира преку отстранување на еден или повеќе жетони од сите негови влезни места, и преку додавање на еден или повеќе жетони во сите негови излезни места, во насока на лаците. После активирањето на преминот, Петриевата мрежа преминува од едно во друго означување. Сите можни означувања во мрежата, добиени при т.н. “игра” со жетони, го формираат множеството на достижност на таа Петриева мрежа, кое може да се претстави графички во облик на граф на достижност. Визуелно тоа е прикажано преку примерот на Слика 1, каде (десно) после

активирање на овозможената транзиција  $t$ , од влезното место  $P_0$  се отстранети двата жетони, а од  $P_1$  е отстранет еден жетон, при што два жетони се сместени во излезното место  $P_2$ , што е во релација со поставените тежини на лаците. Множеството на достижност се состои од иницијалното означување  $(2, 2, 0)$  и следното означување  $(0, 1, 2)$  во кое е донесена Петриевата мрежа после активирањето на преминот  $t$ .



**Слика 1** Пример за динамичкото однесување на едноставна Петриева мрежа, пред (лево) и после (десно) активирање на транзицијата  $t$

Неколку важни карактеристики на Петриевите мрежи, се [45], [50]: живост (liveness) and безбедност (safeness). Една премин се смета за жив ако е потенцијално спремен за активирање. Спротивно на ова, преминот е мртов ако не е потенцијално спремен за активирање. Ако не постои премин која може да се активира, тогаш се вели дека мрежата е во т.н. deadlock ситуација. Едно место е врзано (bounded) ако бројот на негови жетони никогаш не е поголем од некоја позитивна гранична вредност. Исто така, местото е безбедно ако бројот на негови жетони никогаш не е поголем од еден. Во контекст на ова, мрежата е безбедна или врзана ако секое нејзино место е врзано или безбедно, соодветно. Исто така, меѓу најчестите ситуации кои можат да произлезат во една Петриева мрежа [50] се и: секвенцијално извршување и конфликт. Првата ситуација настанува во момент кога еден премин се активира само после активирањето на друг премин. Втората ситуација се појавува во момент кога два или повеќе премини се овозможени заедно (во исто време) во исто означување, така што ако еден од тие премини се активира, останатите ја губат можноста за активација, бидејќи жетонот е веќе отстранет од влезното место.

## 2.1 Генерализирани Стохастички Петриеви мрежи

Однесувањето во стандардните Петриеви мрежи е само дискретно. Тоа значи дека сите премини се моментални, без одлагање, односно моментално се активираат. Една проширена варијанта на овој концепт се Генерализираните Стохастички Петриеви мрежи [1], [45], каде се воведува терминот време и два типа на премини: моментални и временски. Моменталните премини се однесуваат на настани на кои не им е потребно време за активирање. Додека пак временските премини се активираат со одложувања, така што тие одложувања се стохастички, најчесто експоненцијално распределени случајни променливи. Тоа значи дека е поставен тајмер на секој овозможен временски премин, така што вредноста на тајмерот претставува примерок од (негативна) експоненцијална распределба со соодветен интензитет. Тајмерот константно се декрементира, и кога неговата вредност ќе достигне нула, временскиот премин ќе се активира. Најчесто, при графичка репрезентација, моменталните премини се претставени како црни правоаголници или “барови”, додека пак временските премини се прикажани како бели правоаголници. Исто така, моменталните премини се активираат со приоритет врз врменските премини.

Формалната дефиниција на овие мрежи укажува дека секоја GSPN мрежа е составена од следните компоненти [1]:  $PN = (P, T, P, F, H, M_0, W)$ , така што:

1. Првите шест компоненти се однесуваат на стандардната Петриева мрежа каде:
  - a.  $H \subset P \times T$  претставува множество инхибитор лаци (во насока од место до премин) што укажуваат дека преминот не е овозможен во дадената состојба, односно означување,
  - b.  $P$  се однесува на приоритетот на премините, при што временските премини имаат најмал приоритет (0) во однос на моменталните премини чиј приоритет е  $(\geq 1)$  и
2. Последната компонента  $W = (w_1, w_2, \dots, w_n)$  претставува интензитет на одложување на активацијата на преминот доколку преминот е временски, или претставува тежина која се користи за пресметување на веројатноста на активација на преминот доколку преминот е моментален.

Означувањата на мрежата можат да бидат допирливи (tangible) и исчезнувачки (vanishing). Првиот тип на означување постои во момент кога се овозможени само временски премини. Спротивно на ова, вториот тип на означување постои во момент кога се овозможени само моментални премини или комбинација од моментални и временски премини.

## 2.2 Марков Ланец во Континуирано Време

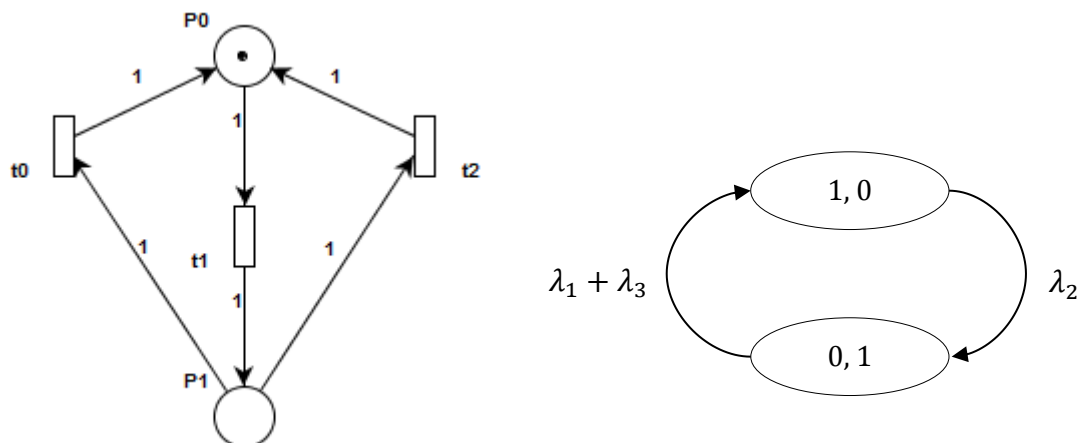
Бидејќи одложувањата на временските премини се стохастички, односно експоненцијално распределени случајни променливи, кои се карактеризираат со т.н. “својство на заоравање”, може да се каже дека концептот на Марков Ланец може да послужи за проучување на однесувањето на еден ваков систем. Тоа својство го подразбира следното: состојбата на системот во време  $t + 1$  зависи само од состојбата на системот во време  $t$ , односно минатото не служи за предвидување на иднината, гледајќи ја сегашноста, прикажано во формула ( 2.1 ):

$$P(X_{t+1} = x_{t+1} | X_t = x_t) \quad (2.1)$$

Преминот на системот од една во друга состојба, односно означување, само помеѓу две времиња  $t$  и  $t - 1$  се опишува преку транзициски веројатности, кои токму поради тоа се познати како едно-чекорни транзициски веројатности. Ваквите транзициски веројатности кои секогаш се ненегативни вредности, се применуваат за конструирање на т.н. транзициска матрица, чија сума по редици е 1, прикажано во формула ( 2.2 ):

$$\sum p_{ij} = 1, \text{ за секое } i \quad (2.2)$$

Множеството состојби во Марковиот Ланец, слично како кај Петриевите мрежи може да се класифицира како затворено ако системот е во една од состојбите на множеството и ќе остане во тоа множество бесконечно. Исто така, доколку множеството состојби содржи состојба со транзициска веројатност  $p_{ii} = 1$ , што значи дека системот бесконечно ќе остане во таа состојба, односно невозможно е да ја напушти таа состојба, се вели дека множеството е абсорбирачко, односно состојбата е абсорбирачка. Останатите состојби се познати како транзитни, односно преодни. Исто така, може да се каже дека графот на достижност кој се генерира за GSPN мрежата, е изоморфен со дијаграмот на интензитетите на состојбите на премин за Марковиот Ланец. Тоа значи, ако означувањата на мрежата се  $\{M_0, M_1, \dots, M_n\}$  тогаш состојбите на Марковиот Ланец се  $\{x_0, x_1, \dots, x_n\}$ . Интензитетот на премин од едено во друго означување, односно од една во друга состојба, се пресметува како сума од интензитетите на активирање на премините кои се овозможени во тоа означување и го носат системот во друго означување, односно состојба. Тоа визуелно може да се види на примерот на Слика 2, каде е илустрирана GSPN мрежа со две места (P0 и P1) и три премини ( $t_0, t_1, t_2$ ), така што секој премин има соодветен интензитет на активација ( $\lambda_1, \lambda_2, \lambda_3$ ). Во ваквата мрежа постојат две означувања {иницијалното  $M_0 (1, 0)$  и  $M_1 (0, 1)$ }, кои соодветствуваат на состојбите во Марковиот Ланец  $\{x_0$  и  $x_1\}$ .



**Слика 2** Пример за GSPN мрежа (лево) и соодветниот дијаграм на интензитетите на состојбите на премин на СТМС (десно)

Сето ова, и останатите детали од [1], укажуваат дека лесно може да се пристапи кон објаснување и анализа на стохастичкиот процес кој стои зад GSPN мрежата со помош на терминологијата на Марков Ланец. Поточно, доколку времето во GSPN мрежата е непрекинато, за опишување се користи концептот на Марков Ланец во Континуирано Време [44].

### 2.3 Користење на GSPN мрежи за моделирање и евалуација на перформанси

Процесот на моделирањето со примена на GSPN мрежа го подразбира следното [1], [45]: конструирање, валидирање и анализа на мрежата, дефинирање на мерки на перформанси, претворање во Марков Ланец и решавање на истиот.

Анализата на GSPN мрежата подразбира: од една страна конструирање на множество на достижност и генерирање граф на достижност, а од друга страна конструирање на матрици. Графот на достижност се генерира според алгоритмот прикажан во [50]:

1. Именувај го иницијалното означување  $M_0$  како корен изначи го како “ново”.
2. За секое ново означување  $M$ :
  - a. Ако  $M$  е идентично означување со друго претходно означување од дрвото,значи го  $M$  како “старо” и премини на друго означување.
  - b. Ако нема транзиции кои се овозможени во  $M$ ,значи го  $M$  како “мртов-крај (dead-end)” и премини на друго означување.
  - c. Се додека постојат овозможени премини во  $M$ ,направи го следното за секој овозможен премин  $t$  во  $M$ :
    - i. Земи го означувањето  $M'$  кое е резултат на активирањето на  $t$  од  $M$ .
    - ii. На патеката од коренот до  $M$ ,ако постои означување  $M''$  такво што  $M'(p) \geq M''(p)$  за секое место  $p$  и  $M' \neq M''$ ,тогаш замени го  $M'$  со  $\omega$  за секое  $p$  такво што  $M'(p) > M''(p)$ .
    - iii. Воведи го  $M'$  како тема,нацртај лак со наслов – од  $M$  до  $Mt'$  и

означи го  $M'$  како “ново”.

3. На крај, спои ги исите темиња од резултатот во граф на достижност.

### Код 1 Алгоритам за креирање граф на достижност

За опишување на влезните и излезните функции на моделот се користат матрици, такви што елементите на т.н. forward incidence матрица се однесуваат на тежините на излезните лаци кои ги поврзуваат премините со местата, елементите на т.н. backward incidence матрица се однесуваат на тежините на влезни лаци, кои ги поврзуваат местата со премините, а комбинираната incidence матрица претставува разлика на forward incidence и backward incidence матриците, што е прикажано во формула ( 2.3 ).

$$A = [a_{ij}] \quad (2.3)$$
$$a_{ij} = a_{ij}^+ - a_{ij}^-$$

Димензијата на сите овие матрици е  $n_t \times n_p$ , каде  $n_t$  е бројот на премини, а  $n_p$  е бројот на места.

Како што е кажано погоре, динамиката на GSPN моделот може да се опише со помош на терминологијата на СТМС [44]. Во таа насока, за стохастичкиот процес кој се наоѓа во основата на GSPN моделот и пресметката на соодветните мерки на перформанси, се користи ефикасен алгоритам во време и простор за пресметување на стабилните состојби на детерминистичките и стохастичките Петриеви мрежи, предложен од [17] и [19]. Алгоритамот претставува анализа на стабилни состојби и пресметка на мерки на перформанси на GSPN моделот. Посебен акцент во првиот дел е ставен на:

1. Градење на дијаграм на интензитетите на состојбите на премин
2. Конструирање на infinitesimal generator матрица

Infinitesimal generator матрицата  $Q$  може да биде конструирана со нумерирање на состојбите од соодветниот Марков Ланец. Елементите што лежат надвор од главната дијагонала претставуваат интензитети на премините од состојбата, додека пак елементите што лежат на главната дијагонала се употребуваат за добивање на сума по редови еднаква на нула, како што е прикажано во формула ( 2.4 ):

$$Q = [q_{ij}]$$
$$q_{ij} = \sum_{t_k \in T_{ij}} \lambda_k \quad (2.4)$$
$$q_{ii} = - \sum_{t_k \in E(M_i)} \lambda_k$$

Во оваа равенка,  $E(M_i)$  се однесува на множеството премини овозможени во означување  $M_i$ ,  $T_{ij}$  е множество премини овозможени во означување  $M_i$  чии активирања го носат системот во друга состојба  $M_j$ , а  $\lambda_k$  е интензитет на активација на преминот  $t_k$ .

3. Решавање на веројатносниот вектор на стационарна состојба

Доколку GSPN мрежата е ергодична<sup>5</sup>, векторот  $\vec{\pi}$  претставува веројатносна распределба на стабилните состојби во дадено означување и може да се пресмета преку решавање на системот линеарни равенки прикажан во формула ( 2.5 ):

$$\begin{cases} \vec{\pi}Q = \vec{0} \\ \sum_i \pi_i = 1 \end{cases} \quad (2.5)$$

Во овој случај,  $\vec{0}$  е вектор со иста димензија како векторот  $\vec{\pi}$ . Имајќи веројатносна распределба, втората равенка се однесува на нормализација (normalization condition).

#### 4. Пресметување на просечното време на престој (average sojourn time)

Просечното време на престој во дадено означување претставува реципрочна вредност од сумата на интензитетите на премините овозможени во тоа означување што ја водат GSPN мрежата од тоа означување во друго означување и се пресметува според формула ( 2.6 ).

$$ST_i = \frac{1}{\sum_{t_j \in E(M_i)} \lambda_j} \quad (2.6)$$

Во оваа равенка, повторно  $E(M_i)$  се однесува на множеството премини овозможени во означување  $M_i$ , а  $\lambda_j$  соодветствува на нивните експоненцијално распределени интензитети на активација.

Неколку мерки на перформанси се користат за евалуација на однесувањето на GSPN моделот. Тие може да бидат пресметани како што е опишано подолу, според [17] и [19].

#### 5. Веројатност на даден настан и просечен број на жетони

Веројатноста на даден настан се пресметува како сума на веројатностите од сите означувања за кои условот, односно настанот е вистинит. Во формулата ( 2.7 ),  $A$  се однесува на множеството означувања за кои условот, односно настанот  $C$  е вистинит.

$$P(C) = \sum_{M_i \in A} \pi_i \quad (2.7)$$

Ако условот, односно настанот се однесува на бројот на жетони во дадено место, просечниот број на жетони во тоа место, уште познат како очекуван број на жетони во тоа место, се пресметува со истата формула погоре.

#### 6. Очекувано време поминато во преодно допирливо означување

Очекувано време поминато во преодни (неапсорбирачки) допирливи означувања пред апсорбирање, се пресметува како во формула ( 2.8 ):

$$\vec{x}Q^N = -\vec{\pi}(0)^N \quad (2.8)$$

<sup>5</sup> [https://en.wikipedia.org/wiki/Markov\\_chain](https://en.wikipedia.org/wiki/Markov_chain)

Тука,  $Q^N$  and  $\vec{\pi}(0)^N$  се ограничувања на infinitesimal generator матрицата  $Q$  и иницијалниот веројатносен вектор на стационарна состојба  $\vec{\pi}(0)$  соодветно, од множеството преодни означувања  $N$ .

7. Кумулативно време на престој во преодно допирливо означување

И ова претставува една интересна мерка која може да биде пресметана за преодните означувања. За да се реализира тоа, за секое означување треба да биде пресметан просечниот број на посети и таа вредност да се множи со соодветното време на престој, како што е прикажано во формула ( 2.9 ):

$$\sigma_i = n_{Ei} * ST_i \quad (2.9)$$

8. Вкупен број на посети во преодно допирливо означување

Очекуваниот вкупен број на посети во преодни состојби, може да се пресмета како во формула ( 2.10 ):

$$n_E(I - P_E^*) = \pi_E^*(0) \quad (2.10)$$

Тука,  $P_E^*$  и  $\pi_E^*$  се ограничувања на веројатносната транзициска матрица  $P_E$  и иницијалниот веројатносен вектор на стационарна состојба  $\vec{\pi}(0)$  соодветно, од множеството преодни означувања  $E$ .

## 2.4 Завршен преглед на поглавјето

Во ова поглавје, беше претставен концептот на Петриеви мрежи, со посебен акцент на Генерализираните Стохастички Петриеви мрежи, преку објаснување на нивните елементи, карактеристики и динамичката природа. Исто така, илустрирано е како концептот на Марков Ланец во Континуирано Време соодветствува на Генерализираните Стохастички Петриеви Мрежи и истиот се применува за нивна анализа и решавање. На крај, наведени се мерките за перформанси кои се користат за евалуација на Генерализираните Стохастички Петриеви мрежи и како истите се пресметуваат со помош на ефикасен алгоритам во време и простор. Во следното поглавје, ќе биде претставен дел од релевантната литература во која опишаните истражувања и стекнати резултати имаат големо значење за нашето истражување.



## ***Глава 3***

### 3 Актуелност на истражувањето

---

Во ова поглавје ќе биде претставен дел од релевантната литература од четири засебни но сепак меѓусебно поврзани подобласти, која беше инспирација за отпочнување на истражување од овој тип и исклучително добро нè воведе и запозна со оваа област.

Во таа насока, да се изгради ефективен модел на однесување на корисникот кој реално ќе го отсликува процесот на пребарување или навигација, значи да се развие прецизен математички модел за предикција. Најчесто, моделите се засноваат на анализа на лог (log) податоци, собрани во одреден временски период [50], [63]. Таквите датотеки се долги и комплексни, но истовремено претставуваат важен извор на информации за корисничкото однесување. Со цел да се анализираат ваквите датотеки и да се извлече одредено знаење за тоа како корисникот пребарува или прави интеракција со веб сајтот, потребана е статистичка анализа и/или примена на техники за податочно рударење. Најчесто, акцентот е ставен на развој на модели за: откривање на шемите на пребарување или навигација [72], предвидување на следни кориснички акции [20] и персонализација врз основа на корисничкото однесување [6], [7], [26], [33], [34], [36], [39], [53], [62], [65], [66], [67], [74], [83].

#### 3.1 Моделирање на однесувањето на корисниците при пребарување

Најчесто користени модели за оваа цел се: Марковите модели од прв ред (First order Markov Models), Маркови модели од повисок ред (High-order Markov Models) и нивните проширувања. На пример, во истражувачката работа на [9], се предлагаат три модели. Првиот модел се однесува на бројот на очекувани кликувања во рамките на една сесија, каде бројот на поставени прашања е однапред познат. Вториот модел се однесува на премините во рамки на една сесија на прашање (query session), за кој се пресметува Марков модел. Во третиот модел пресметуваат временски дистрибуиран транзициски модел, земајќи ги во предвид времето помеѓу поставувањето на прашањето и изборот на релевантен документ. Нивната работа е заснована на лог податоци од Чилеански пребарувач. Нивните резултати укажуваат дека корисниците поставуваат кратки прашања, ги поставуваат само еднаш, избираат неколку страници (документи) и голем дел од нив го реформулираат претходно поставеното прашање со цел да добијат што порелевантни резултати.

[80] користат Скриени Маркови модели (НММ) за моделирање на корисничкото однесување. Нивните експерименти се засноваат на лог податоци собрани во текот на еден ден од комерцијален веб сајт во Холандија, со цел да ги кластерираат корисниците според нивните шеми на однесување при пребарување.

Исто така, интересен аспект е да се моделира промената на однесувањето на корисниците низ времето [57]. Тие применуваат концепти од физика и процесирање на сигнали со цел да предвидат идни промени на корисничкото однесување при пребарување, засновано на историски податоци. Фокусот го имаат ставено на моделирање на промените на фреквенцијата на прашања, кликнати (избрани) веб страни (URL) и парови прашање-URL.

Концептот на Петриеви мрежи како добро развиен графички и математички модел е предложен од [45] за моделирање на однесувањето на корисниците. Покасно, [31] презентираат, но многу површно, концептуална теоретска рамка со чија помош различните кориснички акции во процесот на пребарување може да се опишат како премини.

### **3.2 Моделирање на реформулирањето на прашањето при пребарување**

[58] анализираат реформулирање на прашањата во комбинација со релевантен фидбек (relevance feedback). Меѓу другите анализи, тие испитуваат како се модофицира прашањето во контекст на зголемување и намалување на зборови во него, од едно до друго корисничко прашање и како последователните прашања се разликуваат од останатите прашања од истиот корисник во рамки на иста сесија, претставени како шеми и премини, вклучувајќи ги следните типови на прашања: уникатни, модифицирани, исто прашање - следна страна на резултати и релевантен фидбек.

[70] поставуваат модел на однесување на корисниците така што развиваат три класи на реформулирање на прашањето: содржина, формат и ресурс. Исто така, дефинираат неколку шеми на однесување: специфични, генерализирани, паралелни, блок, динамички, мулти таскинг, повторувачки и променлив формат.

Во истражувањето на [50] реформирирањето на прашањето во една сесија се состои од следните шеми: ново прашање, додавање на зборови, бришење на зборови, повеќе резултати за истото прашање, враќање на претходно прашање и крај на сесијата.

### **3.3 Моделирање на однесувањето на корисниците при навигација**

[61] користат пристап на податочно рударење со цел да ги откријат најчестите шеми на однесување на корисниците при навигација и најчестите секвенци на премини, врз конкретен случај на лог податоци. Исто така, во предвид го земаат и времето на престој на дадена веб страна.

Исто така, пристап на податочно рударење, но во комбинација со семи-Марков процес во дискретно време, се применува во [29] со цел разбирање и опишување на однесувањето

на корисникот. Тие предлагаат и алгоритам за добивање на матрица со веројатностите на премините. Развиениот модел им користи за подобрување на дизајнот на веб сајтот и секако, евалуација на наговите перформанси.

Во истражувањето на [56] се предлага нов пристап за предвидување на однесувањето на корисникот со цел да се подобрат веб перформансите, врз основа на лог податоци и структурата на веб сајтот. Тие применуваат Петриев мрежи со цел да ја откријата (извлечат) структурата на веб сајтот и предвидат следна корисничка акција (path completion).

Модел заснован на Обоени Петриев мрежи (CPN) [30] за предлагање на следни кориснички акции кои се во интерес на корисникот е предложен од страна на [10]. Предлагањето се заснова на профили на претходни корисници и моменталната корисничка сесија.

Уште една примена на Петриевите мрежи се сретнува во истражувањето на [68], но овој пат акцентот е ставен на моделирање и анализирање на структурата на веб сајтот, така што веб страните се претставуваат со помош на состојби, а премините претставуваат линкови. Решавање на ваквиот модел помага во предвидување на следни кориснички акции при навигација (path complete). Ова истражување еволуира, така што [55] го имплементираат концептот на Стохастички Временски Петриев мрежи (STPN) [11] за моделирање на структурата на веб сајтот и предвидување на следно корисничко однесување.

Корисничките стратегии на навигација прикажани во [14], се карактеризирани како: коинцидентни (случајни), генерално прегледување и пребарување. Уште еднаш, и ова истражување е емпириско, односно заклучоците се извлечени од корисничите лог податоци.

Какво е корисничкото онлајн однесување во големи затворени простори при пребарување и навигација (пр. трговски центри), испитано е од страна на [79], со цел да се подобри корисничкото задоволство при користење на интернетот и пазарувањето.

[42] предлагаат моделирање на однесувањето на електронските купувачи со примена на Петриевите мрежи со стохастичко поведење [1], со цел подобрување на квалитетот на веб сервисите, нивната доверливост, перформанси и достапност.

### 3.4 Мерење на немерливото: ангажираност на корисникот

Иако постои палета дефиниции за тоа на што се однесува поимот ангажираност на корисникот (user engagement) [4], [45], [51] сепак може да се издвои следната дефиниција [35]:

*“Ангажираноста на корисникот се однесува на квалитетот на корисничкото искуство со посебен акцент на позитивните аспекти од интеракцијата, а посебно феноменот - привлекување од технологијата.”*

Во различни истражувања се среќаваат бројни мерки за мерење на ангажираност на корисникот, како еден комплексен процес кој тешко може да се измери и за кој нема една единствена унифицирана мерка, и овие мерки пред се зависат од целите на веб сајтот [35]. Покрај “основните” аналитички односно статистички мерки како: времетраење на сесија, број на посети по сесија и број на сесии/посети по корисник и слични мерки прикажани во рамки на истражувањата на [3], [4], [35], [45], [51], може да се измерат и други пропратни мерки кои заедно го отсликуваат корисничкото однесување и укажуваат на корисничкото искуство и ангажираноста на корисникот.

Како такви, можеме да се издвојат мерките conversion rate<sup>6</sup> и abandonment rate<sup>7</sup>. Првата мерка се однесува на процентот на корисници кои превзеле посакувана акција. Втората мерка е потесно поврзана со сферата на електронската трговија, така што се однесува на процентот на корисници кои не станале купувачи, односно предвреме ја напуштиле својата електронска кошничка, без да купат, односно платат. Исто така, можеме да го издвоиме пристапот на [51], каде ангажираноста на корисникот односно посетителот е дефинирана како функција од неколку индекси:

*“Ангажираноста на посетителот е функција од бројот на кликови ( $C_i$ ), времетраењето на посетата ( $D_i$ ), фреквенцијата со која посетителот се враќа на веб сајтот ( $R_i$ ), лојалноста кон веб сајтот ( $L_i$ ), измерената информираност за брендот ( $B_i$ ), желбата за директен фидбек ( $F_i$ ) и веројатноста дека корисниците ќе се ангажираат за специфични активности на веб сајтот дизајнирани да ја зголемат информираноста и да остават трајна импресија ( $I_i$ ).”*

<sup>6</sup> <http://www.nngroup.com/articles/conversion-rates/>

<sup>7</sup> [http://en.wikipedia.org/wiki/Abandonment\\_rate](http://en.wikipedia.org/wiki/Abandonment_rate)

### 3.5 Кластерирање на секвенци од премини

Постојат бројни истражувања за кластерирање на корисничкото однесување. На пример, во трудот на [71] илустриран е нов пристап за кластерирање на шемите на интереси на корисниците, врз основа на регистрираните навигациски податоци од кинеска електронска продавница. Поточно, во предвид освен навигациските патеки, земено се и фреквенцијата на посети на страните и времето посветено, односно времето на задржување на посетителот на дадена категорија.

[72] предлагаат еден нов вид на т.н. хиерархиско кластерирање на корисниците, но истовремено и формирање на содржината на кластерот. Од овој алгоритам произлегува и техника на класификација која лесно може да се искористи за класификација на идни корисници.

Дел од истражувањата во оваа област се насочени кон изградба и имплементација на алгоритми за кластерирање на секвенци од премини, каде не само премините туку и нивниот редослед е важен и носи корисна информација за однесувањето на корисниците. Едно такво истражување е [43], каде се предлага нова архитектура за податочно рударење, заснована врз две техники (алгоритми): асоцијативни правила (association rules) и секвенцијални шеми. Со вакавиот систем се утврдуваат “кориснички трансакции”, што соодветствува на множество кориснички посети, се креира алгоритам за кластерирање на таквите “трансакции”, се применуваат двата алгоритми за податочно рударење и на крај, се прави евалуација на системот. Асоцијативните правила се применети врз “трансакции” неводејќи внимание за времето на нивното случување, а спротивно на ова, се секвенцијалните шеми каде освен за времето на “трансакциите” се води сметка и за поединечните корисници и нивните посети на страни.

Во трудот на [12], развиен е нов метод – Sequence Alignment Method (SAM), кој тестиран врз лог податоци формира неколку кластери според однесувањето на корисниците, односно според редоследот на посетените веб страни и покажува подобри перформанси од стандардните алгоритми за кластерирање кои се засноваат на Евклидово растојание (Euclidean Distance).

Сличен метод за кластерирање на кориснички сесии е промовиран во истражувањето на [77], каде повторно се посветува внимание на редоследот на кориснички акции, а не само нивната содржина. Тие предлагаат нова мерка за сличност заснована на SAM, но со примена на динамичко програмирање.

### 3.6 Завршен преглед на поглавјето

Во ова поглавје прикажан е голем дел од минатите истражувања и нивните заклучоци, во четири меѓусебно поврзани области: моделирање на однесувањето на корисниците при пребарување, моделирање на однесувањето на корисниците при реформулирање на прашање како дел од процесот на пребарување, моделирање на однесувањето на корисниците при навигација, мерење на ангажираност на корисникот и користење на различни техники за кластерирање на секвенци од премини при навигација.

Иако се најде на разноврсни истражувања кои обработуваат различни аспекти од проблемот од интерес, сепак не беа сретнати бројни истражувања кои го претставуваат корисничкото однесување при пребарување или навигација како премини, односно не го применуваат концептот на Петриеви мрежи за моделирање на корисничкото однесување. Токму овој недостаток претставува главна инспирација за оваа докторска дисертација, така што се зачна истражување со цел да испитаме дали Петриевите мрежи со стохастичко поведење можат да се применат за моделирање и опишување на однесувањето на корисниците при пребарување и навигација. Од овој аспект, може да се каже дека оваа дисертација претставува обид да се направи детален приказ и анализа на откриените шеми со примена на Генерелизираните Стохастички Петриеви мрежи, што исто така подразбира кластерирање на корисниците со адаптација на еден од предложените алгоритми кон потребите на истражувањето и исто така претставува прв обид за пресметка на ангажираноста на корисникот преку стохастички модел. Сето ова во детали, преку експерименти и резултати ќе биде опишано во двете следни поглавје.

## ***Глава 4***



## 4 Предложени модели

---

Во ова поглавје ќе бидат претставени предложените бихејвиористички-базирани модели за отсликување на однесувањето на корисниците при процесите на пребарување и навигација, преку примерот на електронска книжарница како информациска мрежа.

### 4.1 Користење на GSPN мрежи за отсликување на шемите на однесување при пребарување

Оваа студија наменета за примената на Петриевите мрежи со стохастичко поведење за отсликување на шемите на корисничко однесување при процесот на пребарување, е заснована на претходна истражувачка работа [50], каде е направен детален приказ на многубројните мерки кои служат за опишување и евалуација на ефикасноста и ефективното на large-scale сервисите за пребарување. Нивните анализи се направени врз голема и комплексна AOL<sup>8</sup> лог датотека со прашања, така што се откриени и опишани шемите на реформулирање на прашање како дел од процесот на пребарување во рамки на една корисничка сесија. Во вака откриените шеми, постојат неколку акции:

- Формулирање на ново прашање (query)
- Модифицирање на дадено прашање: додавање, бришење и промена на зборови
- Враќање на претходно прашање
- Повеќе резултати за дадено прашање (следна страна со резултати)
- Крај на сесијата

Врз основа на секојдневното искуство со пребарување на интернет, воведуваме две дополнителни акции:

- Преглед на резултати кои се вратени
- Избор на даден резултат за понатамошен преглед

Во насока на примена на GSPN мрежите за бихејвиористички-базирано моделирање на процесот на пребарување, развиени се два модели со цел да се прикажат шемите на корисничко однесување во рамки на процесот на реформулација на корисничкото прашање: едноставен и проширен модел. Во GSPN нотација, двете развиени мрежи содржат три места, и неколку временски премини. Детална спецификација е дадена во Табела 1 и Табела 2.

---

<sup>8</sup> <http://www.aol.com/>

**Табела 1 Места во GSPN мрежата**

Place name	Place description	Initial marking
Query	Presence of a query	1
SR	Presence of a list of results	0
ES	Ended user session	0

**Табела 2 Премини во GSPN мрежата**

Transition name	Transition description	Type	Rate
tSR	See results for a given query	Timed	$\alpha$
tSMR	See more results for a given query	Timed	$\lambda$
tAWQ	Add words to the given query	Timed	$\mu$
tCWQ	Change words in the given query	Timed	$\kappa$
tDWQ	Delete words from the given query	Timed	$\nu$
tFNQ	Formulate a new query	Timed	$\theta$
tES	End the session	Timed	$\varepsilon$
tRPQ	Return to a previously asked query	Timed	$\gamma$
tMQ	Modify the query	Timed	$\delta$
tDS	Select a document	Timed	$\beta$

Едноставниот модел е прикажан на Слика 3. Се смета дека сесијата е веќе започната и еден жетон е сместен во влезното место Query, што индицира дека прашањето е веќе поставено. После негативно експоненцијално дистрибуирано време со интензитет  $\alpha$ , преминот tSR се активира, што значи дека пребарување за поставеното прашање е реализирано. Со оваа активација, жетонот е отстранет од влезното место Query и е поставен во излезното место SR. Во овој момент, шест премини се овозможени: tES, tDS, tFNQ, tRPQ, tMQ, и tSMR. Веројатностите за нивно активирање се засновани на нивните активациски интензитети. Секој премин ќе се активира со следната веројатност, соодветно:

$$P(tSMR) = \frac{\lambda}{(\lambda + \delta + \varepsilon + \beta + \theta + \gamma)} \quad (4.1)$$

$$P(tMQ) = \frac{\delta}{(\lambda + \delta + \varepsilon + \beta + \theta + \gamma)} \quad (4.2)$$

$$P(tES) = \frac{\varepsilon}{(\lambda + \delta + \varepsilon + \beta + \theta + \gamma)} \quad (4.3)$$

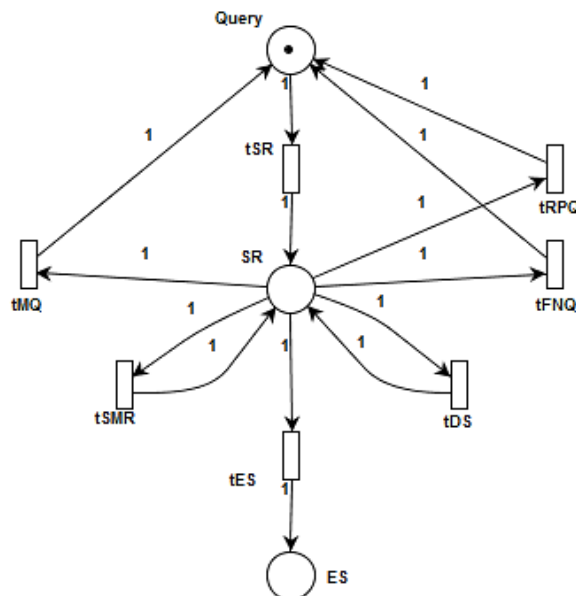
$$P(tRPQ) = \frac{\gamma}{(\lambda + \delta + \varepsilon + \beta + \theta + \gamma)} \quad (4.4)$$

$$P(tFNQ) = \frac{\theta}{(\lambda + \delta + \varepsilon + \beta + \theta + \gamma)} \quad (4.5)$$

$$P(tDS) = \frac{\beta}{(\lambda + \delta + \varepsilon + \beta + \theta + \gamma)} \quad (4.6)$$

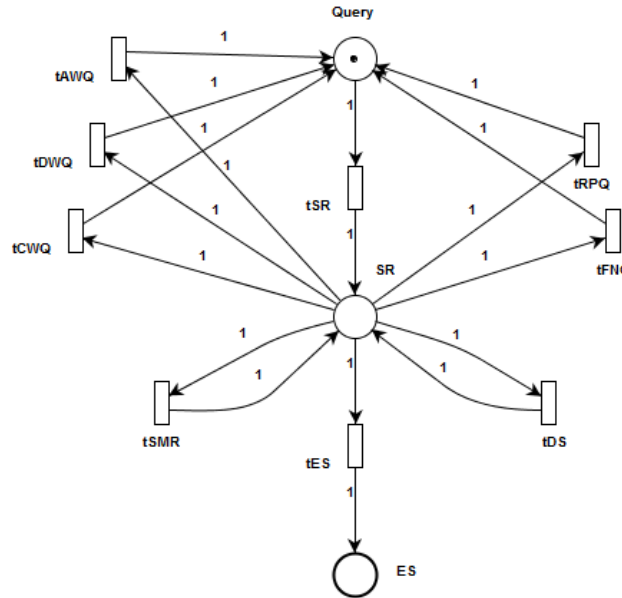
Временските одложувања на овие премини имаат негативна експоненцијална распределба со интензитет:  $\lambda, \delta, \varepsilon, \gamma, \theta, \beta$ , соодветно.

Независно кој премин ќе се активира, жетонот од влезното место SR е отстранет и се појавува во излезното место: ES ако сесијата завршува; SR ако се гледаат повеќе резултати за истото прашање или се избира резултат за понатамошен преглед; Query ако се реализира пребарување за некое претходно прашање или прашањето е модифицирано или ново прашање е поставено. Во последните три ситуации системот е доведен во иницијалната состојба, односно означување, каде временскиот премин tSR е овозможен и се активира со временско одложување со негативна експоненцијална распределба со интензитет  $\alpha$ , преместувајќи го жетонот од влезното место Query во излезното место SR. Ова е повторувачки процес, сè додека сесијата не заврши.



Слика 3 Едноставен модел на корисничко однесување при пребарување

Вториот модел на корисничко однесување се проширува и е прикажан на Слика 4. Сите акции се исти како во едноставниот модел, со исклучок на модификацијата на прашање, која е заменета со три откриени шеми на однесување: додавање на збор/ови, бришење на збор/ови и промена на збор/ови на постоечко прашање.



**Слика 4** Проширен модел на корисничко однесување при пребарување

И во овој случај, сметаме дека сесијата е веќе започната и еден жетон е сместен во влезното место Query, што индицира дека прашањето е веќе поставено. После негативно експоненцијално дистрибуирано време со интензитет  $\alpha$ , преминот tSR се активира, што значи дека пребарување за поставеното прашање е реализирано. Со оваа активација, жетонот е отстранет од влезното место Query и е поставен во излезното место SR. Во овој момент, осум премини се овозможени: tES, tDS, tFNQ, tRPQ, tAWQ, tDWQ, tCWQ и tSMR. Веројатностите за нивно активирање се засновани на нивните брзини. Секој премин ќе се активира со следната веројатност, соодветно:

$$P(tSMR) = \frac{\lambda}{(\lambda + \mu + \nu + \kappa + \varepsilon + \beta + \theta + \gamma)} \quad (4.7)$$

$$P(tAWQ) = \frac{\mu}{(\lambda + \mu + \nu + \kappa + \varepsilon + \beta + \theta + \gamma)} \quad (4.8)$$

$$P(tDWQ) = \frac{\nu}{(\lambda + \mu + \nu + \kappa + \varepsilon + \beta + \theta + \gamma)} \quad (4.9)$$

$$P(tCWQ) = \frac{\kappa}{(\lambda + \mu + \nu + \kappa + \varepsilon + \beta + \theta + \gamma)} \quad (4.10)$$

$$P(tES) = \frac{\varepsilon}{(\lambda + \mu + \nu + \kappa + \varepsilon + \beta + \theta + \gamma)} \quad (4.11)$$

$$P(tRPQ) = \frac{\gamma}{(\lambda + \mu + \nu + \kappa + \varepsilon + \beta + \theta + \gamma)} \quad (4.12)$$

$$P(tFNQ) = \frac{\theta}{(\lambda + \mu + \nu + \kappa + \varepsilon + \beta + \theta + \gamma)} \quad (4.13)$$

$$P(tDS) = \frac{\beta}{(\lambda + \mu + \nu + \kappa + \varepsilon + \beta + \theta + \gamma)} \quad (4.14)$$

Одложувањата на овие премини имаат негативна експоненцијална распределба со интензитет:  $\lambda, \mu, \nu, \kappa, \varepsilon, \gamma, \theta, \beta$  соодветно.

Независно која премин ќе се активира, жетонот е отстранет од влезното место SR и е поставен во едно од следниве влезни места: ES ако сесијата завршува; SR ако се гледаат повеќе резултати за истото прашање или се избира резултат за понатамошен преглед; Query ако се реализира пребарување за некое претходно прашање или прашањето е модифицирано во насока на додавање, бришење или промена на збор/ови или ново прашање е поставено. Во сите последни пет ситуации системот е доведен во иницијалната состојба, односно означување, каде временскиот премин tSR е овозможен и се активира со временско одложување со негативна експоненцијална распределба со интензитет  $\alpha$ , преместувајќи го жетонот од влезното место Query во излезното место SR. Уште еднаш, и овој е повторувачки процес, сè додека сесијата не заврши.

Во продолжение, ќе биде ставен посебен акцент на проширениот модел. Една можна еволуција на овој модел за единечна корисничка сесија е прикажана во Табела 3.

**Табела 3** Можно сценарио на проширениот модел

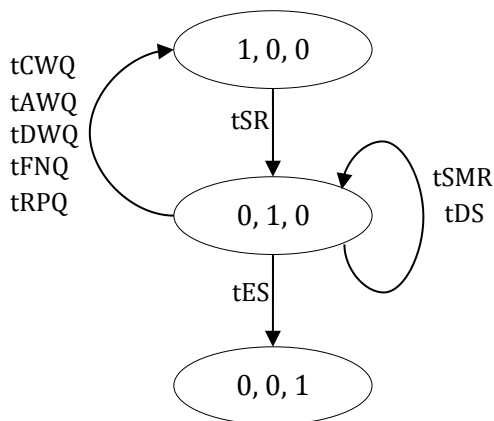
Time	Query History	Query	SR	ES
08:39 am	pattern recognition an algorithmic approach	1	0	0
	<u>tSR</u>	0	1	0
08:39 am	introduction to data mining	1	0	0
	<u>tFNQ</u>	0	1	0
08:45 am	data mining	1	0	0
	<u>tDWQ</u>	0	1	0
08:45 am	select a document	0	1	0
	<u>tDS</u>	0	1	0
08:51 am	pattern recognition an algorithmic approach	1	0	0
	<u>tRPO</u>	0	1	0
08:51 am	select a document	0	1	0
	<u>tDS</u>	0	1	0
08:53 am	select a document	0	1	0
	<u>tDS</u>	0	1	0
09:00 am	pattern recognition	1	0	0
	<u>tDWQ</u>	0	1	0
	tSR	0	1	0

09:02	see more results	0	1	0
am	<u>tSMR</u>			
09:03	End the session	0	0	1
am	<u>tES</u>			

#### 4.1.1 Анализа на предложениот проширен модел

Развиената мрежа е безбедна или врзана бидејќи секое нејзино место е врзано или безбедно, соодветно. Но, во моделот постои и т.н. deadlock ситуација, така што најкратката патека до deadlock е  $tSR \rightarrow tES$ . Исто така, се наидува и на конфликтна ситуација, во моментот кога жетонот е присутен во местото SR, така што сите осум излезни премини  $tES$ ,  $tDS$ ,  $tFNQ$ ,  $tRPQ$ ,  $tAWQ$ ,  $tDWQ$ ,  $tCWQ$  и  $tSMR$  се овозможени во исто време. Со активирањето на било кој еден од овие премини, останатите седум не можат да се активираат. Но со самото активирање на еден премин се разрешува состојбата на конфликт. Исто така овие осум премини можат да се активираат само после активирањето на  $tSR$  преминот, што го воведува моделот во состојба на секвенцијално извршување.

Според [1], [45] врз развиениот модел, реализирана е анализа на достижност со што се откриваат и земаат во предвид сите достапни означувања. За оваа цел, изграден е граф на достижност, прикажан на Слика 5. Иницијалното означување е  $(1, 0, 0)$ , а останатите достапни означувања се  $(0, 1, 0)$  и  $(0, 0, 1)$ . Во овој случај не постојат исчезнувачки означувања. Сите три означувања се допирливи, што го претвора моделот во SPN, но секоја таква мрежа е GSPN.



Слика 5 Граф на достижност

За опишување на влезните и излезните функции на моделот користиме матрици. Според секција 2.3, генерираните матрици се прикажани во Табела 4, Табела 5 и Табела 6, соодветно. Димензијата на сите овие матрици е  $9 \times 3$ , што е еквивалентно на  $n_t \times n_p$ , каде  $n_t$  е бројот на премини, а  $n_p$  е бројот на места.

**Табела 4 Forward incidence матрица I<sup>+</sup>**

	tSR	tSMR	tAWQ	tCWQ	tDWQ	tFNQ	tES	tRPQ	tDS
Query	0	0	1	1	1	1	0	1	0
SR	1	1	0	0	0	0	0	0	1
ES	0	0	0	0	0	0	1	0	0

**Табела 5 Backward incidence матрица I<sup>-</sup>**

	tSR	tSMR	tAWQ	tCWQ	tDWQ	tFNQ	tES	tRPQ	tDS
Query	1	0	0	0	0	0	0	0	0
SR	0	1	1	1	1	1	1	1	1
ES	0	0	0	0	0	0	0	0	0

**Табела 6 Комбинирана incidence матрица I**

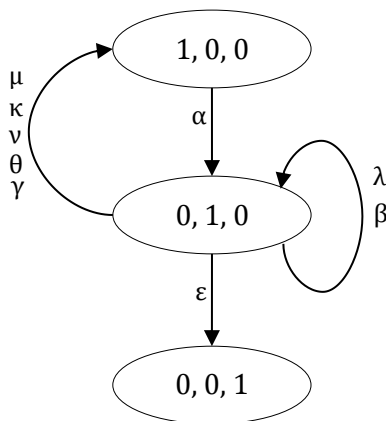
	tSR	tSMR	tAWQ	tCWQ	tDWQ	tFNQ	tES	tRPQ	tDS
Query	-1	0	1	1	1	1	0	1	0
SR	1	0	-1	-1	-1	-1	-1	-1	0
ES	0	0	0	0	0	0	1	0	0

#### 4.1.2 Анализа на стабилни состојби и мерење на перформанси

Сите пресметки во овој дел се реализирани со помош на ефикасен алгоритам во време и простор за пресметување на стабилните состојби на детерминистичките и стохастичките Петриevi мрежи, предложен од [17] и [19], чии чекори се детално опишани во секција 2.3.

##### 1. Градење на дијаграм на интензитетите на состојбите на премин

Соодветниот дијаграм на интензитетите на состојбите на премин на СТМС е прикажан на Слика 6:



**Слика 6 Дијаграм на интензитетите на состојбите на премин на СТМС**

##### 2. Конструирање на infinitesimal generator матрица

Соодветната infinitesimal generator матрица  $Q$ , конструирана со нумерирање на состојбите од соодветниот СТМС е прикажана подолу:

$$Q = \begin{bmatrix} -\alpha & \alpha & 0 \\ x & -x - \varepsilon - y + y & \varepsilon \\ 0 & 0 & 0 \end{bmatrix} \quad (4.15)$$

Каде,

$$\begin{aligned} x &= \mu + \kappa + v + \theta + \gamma \\ y &= \lambda + \beta \end{aligned} \quad (4.16)$$

Во овој случај,  $y$  се игнорира, бидејќи премините tSMR and tDS се премини кон самите себе во означувањето  $(0, 1, 0)$ :

$$Q = \begin{bmatrix} -\alpha & \alpha & 0 \\ x & -x - \varepsilon & \varepsilon \\ 0 & 0 & 0 \end{bmatrix} \quad (4.17)$$

### 3. Решавање на веројатносниот вектор на стационарна состојба

Во соодветниот СТМС, има три состојби за кои трите елементи на векторот  $\vec{\pi}$  се добиваат преку решавање на следниот систем линеарни равенки:

$$\begin{cases} [\pi_0, \pi_1, \pi_2] \begin{bmatrix} -\alpha & \alpha & 0 \\ x & -x - \varepsilon & \varepsilon \\ 0 & 0 & 0 \end{bmatrix} = 0 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{cases} \quad (4.18)$$

Решавање на веројатносниот вектор на стационарна состојба и останатите чекори од алгоритмот кои се однесуваат на пресметката на соодветни мерки за мерење на перформанси, се реализирани во следната секција, врз студја на случај.

#### 4.1.3 Евалуација на мерки на перформанси преку студија на случај

Со цел да се реализира моделирање на однесувањето на корисниците при процесот на пребарување, креирана е лог датотека на клиентска страна, со податоци собирани во текот на осум последователни денови. Станува збор за единечна корисничка сесија, каде корисникот формулира и реформулира серија на прашања. Вклучени се следните информации:

- Query – прашањето за кое се пребарува,
- Query Time – времето кога прашањето е поставено,
- Click URL – ако корисникот кликнал на повратен резултат или избрал документ, неговото URL е зачувано.
- Click Time – времето кога резултатот бил кликнат или документот избран.

Врз основа на статистичката анализа, 62.64% од сите прашања се нови прашања, додека 34.07% се реформулации на постоечки прашања, од кои: 10.99% се прашања на кои се додадени еден или повеќе зборови, 4.40% се прашања во кои е избришан еден или повеќе зборови и 15.68% се прашања со променети зборови. Само 3.30% од прашањата се претходно поставени прашања, а 56.04% од прашањата резултираат со кликање (избор) на еден или повеќе резултати или избор на документи. Интензитет на



формулирање/реформулирање на прашање претставува средна вредност на формулирање/реформулирање на прашање во единица време, прикажано со формула ( 4.19 ).

$$QFRrate = \frac{1}{mean\ QFRT} \quad (4.19)$$

Во претходната формула,  $QFRrate$  претставува интензитет на формулирање/реформулирање на прашање, а  $mean\ QFRT$  е средна вредност на формулирање/реформулирање на прашање во единица време. Во овој случај, времето за кое се прават сите анализи е изразено во секунди.

Според добиените квантитативни податоци, интензитетите на сите девет временски премини во изградената GSPN мрежа, се прикажани во Табела 7.

**Табела 7 Интензитет на временските премини**

	Rate
$\alpha$	3.03030
$\lambda$	3.03030
$\mu$	0.01192
$\kappa$	0.00481
$\nu$	0.00667
$\theta$	0.00011
$\varepsilon$	0.00034
$\gamma$	0.00002
$\beta$	0.03011

Веројатностите за активирање на временските премини се прикажани во Табела 8:

**Табела 8 Веројатности на активација**

	Probability
$P(tSR)$	1.000000
$P(tSMR)$	0.495586
$P(tAWQ)$	0.001949
$P(tDWQ)$	0.001091
$P(tCWQ)$	0.000786
$P(tES)$	0.000055
$P(tRPQ)$	0.000004
$P(tFNQ)$	0.000017
$P(tDS)$	0.004924

Веројатносните дистрибуции на стабилните состојби во соодветниот СТМС, се пресметуваат со примена на следниот систем линеарни равенки, чиј опис е даден во секција 2.3, формула ( 2.5 ):

$$\begin{cases} [\pi_0, \pi_1, \pi_2] \begin{bmatrix} -3.03030 & 3.03030 & 0 \\ 0.02353 & -0.02387 & 0.00034 \\ 0 & 0 & 0 \end{bmatrix} = 0 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{cases} \quad (4.20)$$

Од ова може да се заклучи дека  $\pi_0 = \pi_1 = 0$  и  $\pi_2 = 1$ , бидејќи последната состојба е апсорбирачка, што значи пресметката на веројатносните дистрибуции на стабилните состојби е безначајна. Сите други мерки од интерес се пресметани и прикажани подолу. Просечните времиња на престој за допирливите состојби се прикажани во Табела 9, добиени според формула ( 2.6 ) од секција 2.3:

**Табела 9 Просечно време на престој**

Marking	Time (seconds)
(1, 0, 0)	0.33
(0, 1, 0)	41.89
(0, 0, 1)	$\infty$

Просечното време за престој за првата состојба е 0.33 [s], а за втората состојба е 41.89 [s]. Исто така вкупното време на престој во овие две допирливи состојби може да биде пресметано според формула ( 2.8 ) од секција 2.3:

$$[x_0, x_1] \begin{bmatrix} -3.03030 & 3.03030 \\ 0.02353 & -0.02387 \end{bmatrix} = [-1, 0] \quad (4.21)$$

Од овие равенки,  $x_0 = 23.17$  и  $x_1 = 2941.17$ , што значи дека вкупното време поминато во првата состојба е 23.17 [s] и 2941.17 [s] во втората состојба.

Просечниот број на посети во двете преодни состојби е 70.2 посети. Оваа вредност се добива со помош со следните равенки добиени според формула ( 2.10 ) од секција 2.3:

$$[n_0, n_1] \begin{bmatrix} 1 & -1 \\ -0.98575 & 1 \end{bmatrix} = [1, 0] \quad (4.22)$$

$$n_1 = n_0 = 70.2$$

Во моделот, кумулативното време на престој за првото преодно означување е 23.16 [s] and 2940.68 [s] за второто преодно означување, пресметано според чекор 7 од секција 2.3:

$$\sigma_1 = 70.2 * 0.33 = 23.166$$

$$\sigma_2 = 70.2 * 41.89 = 2940.678 \quad (4.23)$$

Оттука следува дека, просечното време на траење на единечна корисничка сесија е 2963.84 [s] = 49.4 [min].

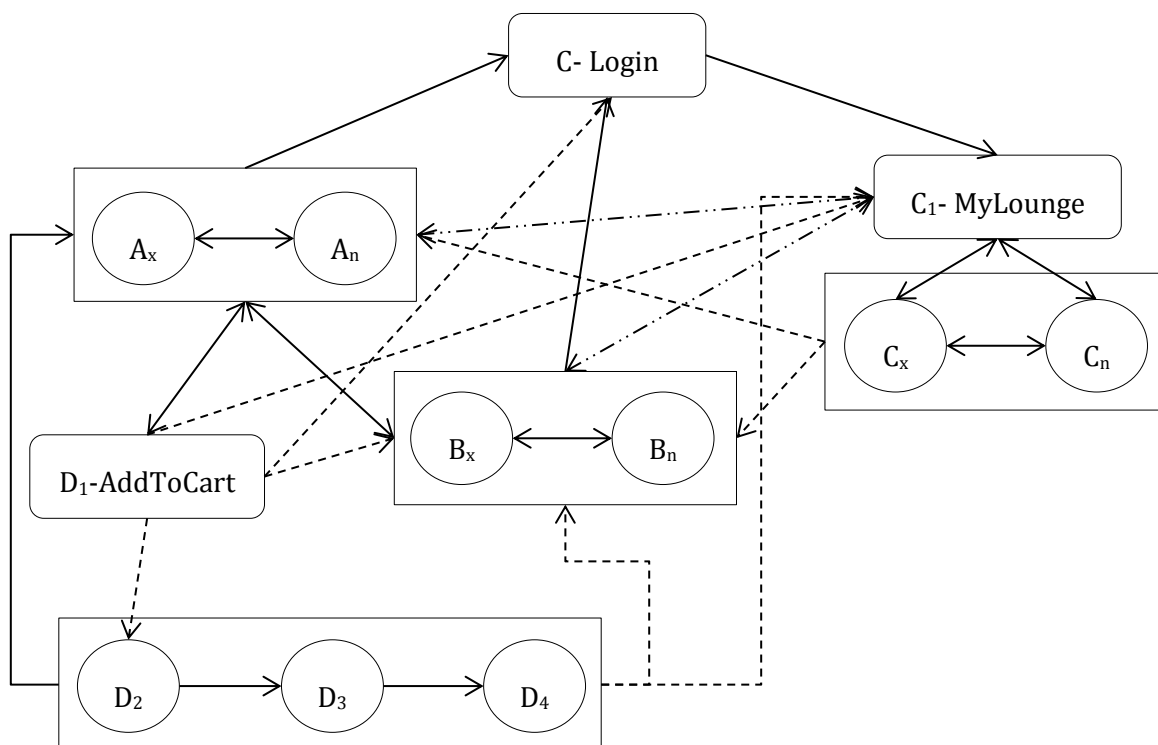
## 4.2 Користење на GSPN мрежи за отсликување на шемите на однесување при навигација

Студијата наменета за примената на Петриевите мрежи со стохастичко поведение за отсликување на шемите на однесување при процесот на навигација, е инспирирана од претходна истражувачка работа на [42], [55], [68], каде Петриевите мрежи со стохастичко поведение се применуваат за моделирање на однесувањето на електронските купувачи и анализа на структурата на веб сајтот и предвидување на следни кориснички акции, соодветно.

Соодветните анализи за откривање на однесувањето на корисниците како и предложениот модел во рамки на оваа докторска дисертација, се засноваат врз навигациски податоци собрани од првата електронска книжарница во Република Македонија [3], но лесно генерализирани и применливи врз познатите сервиси, чија структура на системот е во голема сличност со избраното сценарио.

### 4.2.1 Предложен модел

Електронската книжарница располага со соодветна структура (топологија), која е прикажана на Слика 7.



Слика 7 Структура на веб страната

Истата содржи страни кои се класифицирани како страни од категорија: А, В, С, и D. Во категорија А спаѓаат страни кои се однесуваат на производите (книгите) што се продаваат, како на пример: преглед на книги по категорија, детали за дадена книга, попусти, нови книги, и сл. Во категорија В спаѓаат страни кои се однесуваат на информации за електронската книжарница, како на пример: помош околу купувањето, политика на испорака, заштита на лични податоци и сл. Страните од категорија С се достапни само во случај на логиран корисник и се однесуваат на: податоците за самиот корисник, неговиот профил, неговите нарачки и сл. На крај, страните од категорија D се однесуваат на чекорите на купување и плаќање. Целокупен преглед на страните по категории е даден во Табела 10.

**Табела 10 Листа на страни по категории**

Категорија	Страна (англиски)	Страна (македонски)
А	Registration	Регистрација
	ListBooksforCategory	Преглед на производи (книги) по категорија
	Default	Почетна
	NewProducts	Нови производи (книги)
	BestSellers	Најпродавани производи (книги)
	ProductDiscounts	Попусти
	News	Новости
	Publisher	Издавачи
	BookStoreMenu	Книжарници
	SearchPreview	Пребарување
	ContactUs	Контакт
	Help	Помош
	BookDetails	Детали за даден производ (книга)
ReadPDF	Читање на исечок од книга	
В	AboutUs	За нас
	GeneralProvisions	Провизии
	ProtectPersonalData	Заштита на лични податоци
	DeliveryPolitics	Политика на испорака
	BackPolicy	Политика на враќање
	HelpOrders	Помош при нарачки
	HelpRegistration	Помош при регистрација
Marketing	Маркетинг	
С	LogIn	Логирање
	MyLoungeDef	Моја библиотека
	UserProfile	Кориснички профил
	UserAddress	Корисничка адреса

	UserOrders	Нарачки на корисник
	SavedForLater	Зачувано за подоцна
	NewsLetter	Весник
	UserReferrals	Покани пријател
	ImportCSV	Листа на пријатели кои се покануваат
	HelpUser	Помош за корисник
D	AddToCart	Додади во кошничка
	LastCartPreview	Последен преглед на кошничка
	PayingCasys2	Плаќање
	AddressEntry	Избор (промена) на адреса на достава

---

Во откриените шеми на навигација, според структурата на електронската книжарница, може да се издвојат следните кориснички акции:

- Преглед на страна од категорија А
- Преглед на страна од категорија В
- Крај на посетата
- Логирање на корисник
- Преглед на страната која се однесува на корисничката сметка (библиотека) на веќе логиран корисник
- Преглед на страна од категорија С
- Додавање на производ (книга) во кошничка
- Избор на адреса за достава
- Последен преглед на кошничка
- Купување/Плаќање

И во овој случај, динамичката природа на GSPN мрежите е прилагодлива за моделирање на процесот на навигација. Во оваа насока, развиен е модел кој го детално ги опишува откриените шеми на однесување на корисниците при процесот на навигација во рамки на електронската книжарница. Во GSPN нотација, мрежата содржи 10 места и 44 временски премини. Детална спецификација е дадена во Табела 11 и Табела 12, соодветно.

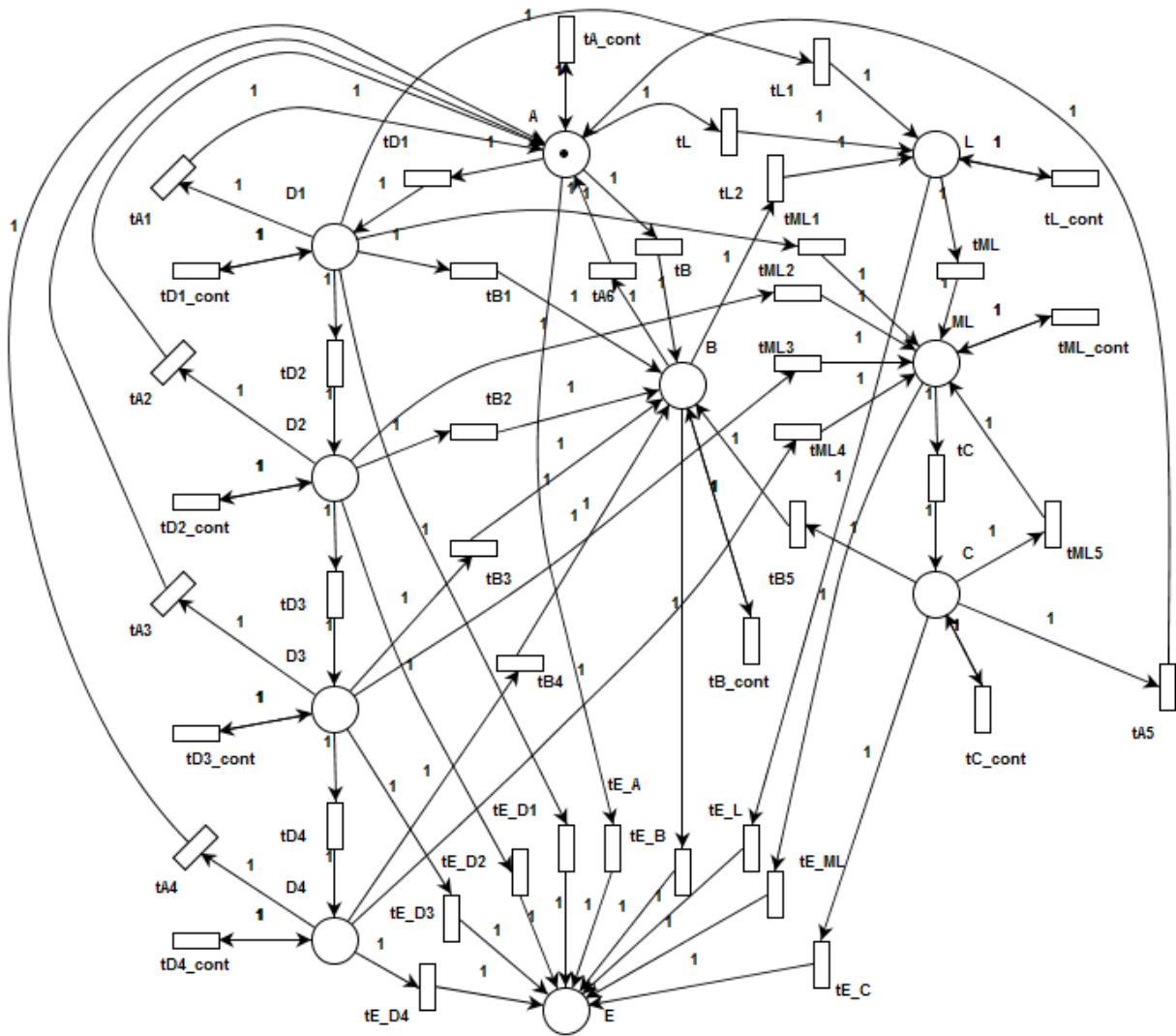
**Табела 11 Места во GSPN мрежата**

Place name	Place description	Initial marking
A	A category page	1
B	B category page	0
E	Ended user session	0
L	Login	0
ML	MyLounge (MyBookStore)	0
C	C category page	0
D1	AddToCart	0
D2	AddressEntry	0
D3	LastCartPreview	0
D4	PayingCasys	0

**Табела 12 Премини во GSPN мрежата**

Transition name	Transition description	Type	Rate
tA_cont, tA1, tA2, tA3, tA4, tA5, tA6	Visit an A category page	Timed	$\alpha$
tB, tB_cont, tB1, tB2, tB3, tB4, tB5	Visit a B category page	Timed	$\lambda$
tE_A, tE_B, tE_L, tE_ML, tE_C, tE_D1, tE_D2, tE_D3, tE_D4	End the session	Timed	$\mu$
tL, tL1, tL2, tL_cont	Login	Timed	$\kappa$
tML, tML1, tML2, tML3, tML4, tML5, tML_cont	Visit MyLounge page	Timed	$\nu$
tC, tC_cont	Visit a C category page	Timed	$\theta$
tD1, tD1_cont	Visit AddToCart page	Timed	$\varepsilon$
tD2, tD2_cont	Visit AddressEntry page	Timed	$\gamma$
tD3, tD3_cont	Visit LastCartPreview page	Timed	$\delta$
tD4, tD4_cont	Visit PayingCasys page	Timed	$\beta$

Графички приказ на моделот е прикажан на Слика 8.



Слика 8 Модел на корисничко однесување при навигација

Во овај случај, се смета дека посетата е веќе започната и еден жетон е сместен во влезното место A, што индицира дека истата веќе започнала со преглед на страна од категорија A. Кога жетонот е во местото A, овозможени се следните премини:  $tA\_cont$ ,  $tB$ ,  $tE\_A$ ,  $tL$  и  $tD1$ . Секој овозможен премин ќе се активира со следната веројатност, соодветно:

$$P(tA\_cont) = \frac{\alpha}{(\alpha + \lambda + \kappa + \varepsilon + \mu)} \quad (4.24)$$

$$P(tB) = \frac{\lambda}{(\alpha + \lambda + \kappa + \varepsilon + \mu)} \quad (4.25)$$

$$P(tE\_A) = \frac{\mu}{(\alpha + \lambda + \kappa + \varepsilon + \mu)} \quad (4.26)$$

$$P(tL) = \frac{\kappa}{(\alpha + \lambda + \kappa + \varepsilon + \mu)} \quad (4.27)$$

$$P(tD1) = \frac{\varepsilon}{(\alpha + \lambda + \kappa + \varepsilon + \mu)} \quad (4.28)$$

После временско одложување со негативна експоненцијална распределба со соодветен интензитет, може да се активира еден од овозможените премини:

- Доколку станува збор за преминот  $tA\_cont$ , интензитетот е  $\alpha$ , така што после активирањето на овој премин, системот е доведен во иницијалната состојба, односно преглед на страна од категорија А после преглед исто така на страна од категорија А.
- Интензитетот кој се однесува на преминот  $tB$  е  $\lambda$ , така што при активирање на овој премин жетонот е преместен од местото А во местото В, што индицира дека корисникот прегледал страна од категорија В после прегледување на страна од категорија А.
- Доколку се активира преминот  $tL$ , интензитетот е  $\kappa$ , така што жетонот е преместен од местото А во местото L, што укажува дека корисникот пристапил кон логирање, после прегледување на страна од категорија А.
- Слична е ситуацијата и со преминот  $tD1$  чиј интензитет е  $\varepsilon$ , но во овој случај, при нејзино активирање жетонот е преместен од местото А во местото D1, што претставува ствање во кошничка да одреден производ после прегледување на страна од категорија А.
- Кога корисникот се одлучил за да ја заврши својата посета после прегледување на страна од категорија А, се активира преминот  $tE\_A$  со интензитет  $\mu$ , така што жетонот е преместен од местото А во местото Е.

Кога жетонот е во местото В, овозможени се премините:  $tA6$ ,  $tB\_cont$ ,  $tE\_B$  и  $tL2$ . Секој премин ќе се активира со следната веројатност, соодветно:

$$P(tA6) = \frac{\alpha}{(\alpha + \lambda + \mu + \kappa)} \quad (4.29)$$

$$P(tB\_cont) = \frac{\lambda}{(\alpha + \lambda + \mu + \kappa)} \quad (4.30)$$

$$P(tE\_B) = \frac{\mu}{(\alpha + \lambda + \mu + \kappa)} \quad (4.31)$$

$$P(tL2) = \frac{\kappa}{(\alpha + \lambda + \mu + \kappa)} \quad (4.32)$$

После временско одложување со негативна експоненцијална распределба со соодветен интензитет, може да се активира еден од овозможените премини:

- Доколку станува збор за преминот  $tA6$ , интензитетот е  $\alpha$ , така што после активирањето на овој премин, системот е доведен во иницијалната состојба, односно прегледување на страна од категорија А после прегледување на страна од категорија В.



- Со активирање на преминот  $t_{V\_cont}$ , интензитетот е  $\lambda$ , така што системот е доведен во истата состојба, односно прегледување настрана од категорија В после прегледување исто така на страна од категорија В.
- Доколку се активира преминот  $t_{L2}$ , интензитетот е  $\kappa$ , така што жетонот е преместен од местото В во местото L, што укажува дека корисникот пристапил кон логирање, после прегледување на страна од категорија В.
- Кога корисникот се одлучил за да ја заврши својата посета после прегледување на страна од категорија В, се активира преминот  $t_{E\_V}$  со интензитет  $\mu$ , така што жетонот е преместен од местото В во местото Е.

Кога жетонот е во местото L, овозможени се премините:  $t_E$ ,  $t_L$  и  $t_{ML}$ . Секој премин ќе се активира со следната веројатност, соодветно:

$$P(t_{E\_L}) = \frac{\mu}{(\mu + \kappa + \nu)} \quad (4.33)$$

$$P(t_{L\_cont}) = \frac{\kappa}{(\mu + \kappa + \nu)} \quad (4.34)$$

$$P(t_{ML}) = \frac{\nu}{(\mu + \kappa + \nu)} \quad (4.35)$$

После временско одложување со негативна експоненцијална распределба со соодветен интензитет, може да се активира еден од овозможените премини:

- Кога се активира преминот  $t_{ML}$ , интензитетот е  $\nu$ , така што после активирањето на овој премин веќе логираниот корисник пристапил кон страната што се однесува на неговиот профил, односно неговата библиотека.
- Со активирање на преминот  $t_{L\_cont}$ , интензитетот е  $\kappa$ , така што системот е доведен во истата состојба, односно повторна прегледување на страната за логирање.
- Кога корисникот се одлучил за да ја заврши својата посета после процесот на логирање, се активира преминот  $t_{E\_L}$  со интензитет  $\mu$ , така што жетонот е преместен од местото L во местото Е.

Кога жетонот е во местото D1, овозможени се премините:  $t_{A1}$ ,  $t_{B1}$ ,  $t_{E\_D1}$ ,  $t_{L1}$ ,  $t_{ML1}$ ,  $t_{D1\_cont}$  и  $t_{D2}$ . Секој премин ќе се активира со следната веројатност, соодветно:

$$P(t_{A1}) = \frac{\alpha}{(\alpha + \lambda + \mu + \kappa + \nu + \varepsilon + \gamma)} \quad (4.36)$$

$$P(t_{B1}) = \frac{\lambda}{(\alpha + \lambda + \mu + \kappa + \nu + \varepsilon + \gamma)} \quad (4.37)$$

$$P(t_{E\_D1}) = \frac{\mu}{(\alpha + \lambda + \mu + \kappa + \nu + \varepsilon + \gamma)} \quad (4.38)$$

$$P(t_{L1}) = \frac{\kappa}{(\alpha + \lambda + \mu + \kappa + \nu + \varepsilon + \gamma)} \quad (4.39)$$

$$P(tML1) = \frac{\nu}{(\alpha + \lambda + \mu + \kappa + \nu + \varepsilon + \gamma)} \quad (4.40)$$

$$P(tD1\_cont) = \frac{\varepsilon}{(\alpha + \lambda + \mu + \kappa + \nu + \varepsilon + \gamma)} \quad (4.41)$$

$$P(tD2) = \frac{\gamma}{(\alpha + \lambda + \mu + \kappa + \nu + \varepsilon + \gamma)} \quad (4.42)$$

После временско одложување со негативна експоненцијална распределба со соодветен интензитет, може да се активира еден од овозможените премини:

- Доколку станува збор за преминот  $tA1$ , интензитетот е  $\alpha$ , така што после активирањето на овој премин, системот е доведен во иницијалната состојба, односно прегледување на страна од категорија А после ставање на производ во својата кошничка.
- Интензитетот кој се однесува на преминот  $tB1$  е  $\lambda$ , така што при активирање на овој премин жетонот е преместен од местото D1 во местото B, што индицира дека корисникот посетил страна од категорија B после ставање на производ во својата кошничка.
- Кога се активира преминот  $tML1$ , интензитетот е  $\nu$ , така што после активирањето на овој премин веќе логираниот корисник пристапил кон страната што се однесува на неговиот профил, односно неговата библиотека, после ставање на производ во својата кошничка, односно жетонот е преместен од местото D1 во местото ML.
- Со активирање на преминот  $tL1$ , интензитетот е  $\kappa$ , така што откако корисникот ставил производ во својата кошничка се одлучил за логирање, најверојатно со цел да може да го продолжи процесот на купување, односно жетонот е преместен од местото D1 во местото L.
- Кога корисникот се одлучил за да ја заврши својата посета после процесот на ставање производ во својата кошничка, се активира преминот  $tE\_D1$  со интензитет  $\mu$ , така што жетонот е преместен од местото D1 во местото E.
- Со повторно активирање на преминот  $tD1\_cont$  со интензитет  $\varepsilon$ , системот е доведен во истата состојба, односно повторен преглед на својата кошничка.
- Кога корисникот се одлучил да премине на следниот чекор во процесот на купување/плаќање, се активира преминот  $tD2$  со интензитет  $\gamma$ , така што жетонот е преместен од местото D1 во местото D2.

Во контекст на продолжување на процесот на купување/плаќање, кога жетонот е во местото D2, овозможена е преминот  $tD3$ , но и премините:  $tA2$ ,  $tB2$ ,  $tE\_D2$ ,  $tML2$  и  $tD2\_cont$ . Секој премин ќе се активира со следната веројатност, соодветно:

$$P(tA2) = \frac{\alpha}{(\alpha + \lambda + \mu + \nu + \gamma + \delta)} \quad (4.43)$$

$$P(tB2) = \frac{\lambda}{(\alpha + \lambda + \mu + \nu + \gamma + \delta)} \quad (4.44)$$

$$P(tE\_D2) = \frac{\mu}{(\alpha + \lambda + \mu + \nu + \gamma + \delta)} \quad (4.45)$$

$$P(tML2) = \frac{\nu}{(\alpha + \lambda + \mu + \nu + \gamma + \delta)} \quad (4.46)$$

$$P(tD2\_cont) = \frac{\gamma}{(\alpha + \lambda + \mu + \nu + \gamma + \delta)} \quad (4.47)$$

$$P(tD3) = \frac{\delta}{(\alpha + \lambda + \mu + \nu + \gamma + \delta)} \quad (4.48)$$

После временско одложување со негативна експоненцијална распределба со соодветен интензитет, може да се активира еден од овозможените премини:

- Доколку станува збор за преминот  $tA2$ , интензитетот е  $\alpha$ , така што после активирањето на овој премин, системот е доведен во иницијалната состојба, односно прегледување на страна од категорија А после промена на адреса на достава на избраниот производ.
- Интензитетот кој се однесува на преминот  $tB2$  е  $\lambda$ , така што при активирање на овој премин жетонот е преместен од местото D2 во местото B, што индицира дека корисникот посетил страна од категорија B после промена на адреса на достава на избраниот производ.
- Кога се активира преминот  $tML2$ , интензитетот е  $\nu$ , така што после активирањето на овој премин веќе логираниот корисник пристапил кон страната што се однесува на неговиот профил, односно неговата библиотека, после промена на адреса на достава на избраниот производ, односно жетонот е преместен од местото D2 во местото ML.
- Кога корисникот се одлучил за да ја заврши својата посета после промена на адреса на достава на избраниот производ, се активира преминот  $tE\_D2$  со интензитет  $\mu$ , така што жетонот е преместен од местото D2 во местото E.
- Со повторно активирање на преминот  $tD2\_cont$  со интензитет  $\gamma$ , системот е доведен во истата состојба, односно повторна промена на адресата на достава на избраниот производ.
- Кога корисникот се одлучил да премине на следниот чекор во процесот на купување/плаќање, се активира преминот  $tD3$  со интензитет  $\delta$ , така што жетонот е преместен од местото D2 во местото D3.

Понатаму, кога жетонот е во местото D3, овозможена е преминот tD4, но и премините: tA3, tB3, tE\_D3, tML3 и tD3\_cont, tD4. Секој премин ќе се активира со следната веројатност, соодветно:

$$P(tA3) = \frac{\alpha}{(\alpha + \lambda + \mu + \nu + \delta + \beta)} \quad (4.49)$$

$$P(tB3) = \frac{\lambda}{(\alpha + \lambda + \mu + \nu + \delta + \beta)} \quad (4.50)$$

$$P(tE_D3) = \frac{\mu}{(\alpha + \lambda + \mu + \nu + \delta + \beta)} \quad (4.51)$$

$$P(tML3) = \frac{\nu}{(\alpha + \lambda + \mu + \nu + \delta + \beta)} \quad (4.52)$$

$$P(tD3\_cont) = \frac{\delta}{(\alpha + \lambda + \mu + \nu + \delta + \beta)} \quad (4.53)$$

$$P(tD4) = \frac{\beta}{(\alpha + \lambda + \mu + \nu + \delta + \beta)} \quad (4.54)$$

После временско одложување со негативна експоненцијална распределба со соодветен интензитет, може да се активира еден од овозможените премини:

- Доколку станува збор за преминот tA3, интензитетот е  $\alpha$ , така што после активирањето на овој премин, системот е доведен во иницијалната состојба, односно прегледување на страна од категорија A после последен преглед на својата кошничка.
- Интензитетот кој се однесува на преминот tB3 е  $\lambda$ , така што при активирање на овој премин жетонот е преместен од местото D3 во местото B, што индицира дека корисникот посетил страна од категорија B после последен преглед на својата кошничка.
- Кога се активира преминот tML3, интензитетот е  $\nu$ , така што после активирањето на овој премин веќе логираниот корисник пристапил кон страната што се однесува на неговиот профил, односно неговата библиотека, после последен преглед на својата кошничка, односно жетонот е преместен од местото D3 во местото ML.
- Кога корисникот се одлучил за да ја заврши својата посета после после последен преглед на својата кошничка, се активира преминот tE\_D3 со интензитет  $\mu$ , така што жетонот е преместен од местото D3 во местото E.
- Со повторно активирање на преминот tD3\_cont со интензитет  $\delta$ , системот е доведен во истата состојба, односно повторно за последен пат ја прегледува својата кошничка.
- Кога корисникот се одлучил да премине на последниот чекор во процесот на купување/плаќање, се активира преминот tD4 со интензитет  $\beta$ , така што жетонот е преместен од местото D3 во местото D4.

Кога жетонот е во место D4, овозможени се премините: tA4, tB4, tE\_D4, tML4 и tD4\_cont.

Секој премин ќе се активира со следната веројатност, соодветно:

$$P(tA4) = \frac{\alpha}{(\alpha + \lambda + \mu + \nu + \beta)} \quad (4.55)$$

$$P(tB4) = \frac{\lambda}{(\alpha + \lambda + \mu + \nu + \beta)} \quad (4.56)$$

$$P(tE\_D4) = \frac{\mu}{(\alpha + \lambda + \mu + \nu + \beta)} \quad (4.57)$$

$$P(tML4) = \frac{\nu}{(\alpha + \lambda + \mu + \nu + \beta)} \quad (4.58)$$

$$P(tD4\_cont) = \frac{\beta}{(\alpha + \lambda + \mu + \nu + \beta)} \quad (4.59)$$

После временско одложување со негативна експоненцијална распределба со соодветен интензитет, може да се активира еден од овозможените премини:

- Доколку станува збор за преминот tA4, интензитетот е  $\alpha$ , така што после активирањето на овој премин, системот е доведен во иницијалната состојба, односно прегледување на страна од категорија A после процесот на купување/плаќање.
- Интензитетот кој се однесува на преминот tB4 е  $\lambda$ , така што при активирање на овој премин жетонот е преместен од местото D4 во местото B, што индицира дека корисникот посетил страна од категорија B после процесот на купување/плаќање.
- Кога се активира преминот tML4, интензитетот е  $\nu$ , така што после активирањето на овој премин веќе логираниот корисник пристапил кон страната што се однесува на неговиот профил, односно неговата библиотека, после процесот на купување/плаќање, односно жетонот е преместен од местото D4 во местото ML.
- Кога корисникот се одлучил за да ја заврши својата посета после процесот на купување/плаќање, се активира преминот tE\_D4 со интензитет  $\mu$ , така што жетонот е преместен од местото D4 во местото E.
- Со повторно активирање на преминот tD4\_cont со интензитет  $\beta$ , системот е доведен во истата состојба, односно повторно активирање на процесот на купување/плаќање.

Кога жетонот е во местото ML, овозможени се премините: tE\_ML, tML\_cont и tC. Секој премин ќе се активира со следната веројатност, соодветно:

$$P(tE_{ML}) = \frac{\mu}{(\mu + \nu + \theta)} \quad (4.60)$$

$$P(tML_{cont}) = \frac{\nu}{(\mu + \nu + \theta)} \quad (4.61)$$

$$P(tC) = \frac{\theta}{(\mu + \nu + \theta)} \quad (4.62)$$

После временско одложување со негативна експоненцијална распределба со соодветен интензитет, може да се активира еден од овозможените премини:

- Кога се активира преминот  $tML_{cont}$ , интензитетот е  $\nu$ , така што после активирањето на овој премин логираниот корисник повторно пристапил кон страната што се однесува на неговиот профил, односно неговата библиотека.
- Со активирање на преминот  $tC$ , интензитетот е  $\theta$ , така што жетонот е преместен од местото  $ML$  во местото  $C$ , што значи прегледување на една од страните од категорија  $C$  после прегледување на својата библиотека.
- Кога корисникот се одлучил за да ја заврши својата посета после посетата на својата библиотека, се активира преминот  $tE_{ML}$  со интензитет  $\mu$ , така што жетонот е преместен од местото  $ML$  во местото  $E$ .

Кога жетонот е во местото  $C$ , овозможени се премините:  $tA5$ ,  $tB5$ ,  $tE_C$ ,  $tML5$  и  $tC_{cont}$ . Секој премин ќе се активира со следната веројатност, соодветно:

$$P(tA5) = \frac{\alpha}{(\alpha + \lambda + \mu + \nu + \theta)} \quad (4.63)$$

$$P(tB5) = \frac{\lambda}{(\alpha + \lambda + \mu + \nu + \theta)} \quad (4.64)$$

$$P(tE_C) = \frac{\mu}{(\alpha + \lambda + \mu + \nu + \theta)} \quad (4.65)$$

$$P(tML5) = \frac{\nu}{(\alpha + \lambda + \mu + \nu + \theta)} \quad (4.66)$$

$$P(tC_{cont}) = \frac{\theta}{(\alpha + \lambda + \mu + \nu + \theta)} \quad (4.67)$$

После временско одложување со негативна експоненцијална распределба со соодветен интензитет, може да се активира еден од овозможените премини:

- Доколку станува збор за преминот  $tA5$ , интензитетот е  $\alpha$ , така што после активирањето на овој премин, системот е доведен во иницијалната состојба, односно прегледување на страна од категорија  $A$  после прегледување на страна од категорија  $C$ .
- Интензитетот кој се однесува на преминот  $tB5$  е  $\lambda$ , така што при активирање на овој премин жетонот е преместен од местото  $C$  во местото  $B$ , што индицира дека корисникот посетил страна од категорија  $B$  после прегледување на страна од категорија  $C$ .

- Кога се активира преминот  $t_{ML5}$ , интензитетот е  $\nu$ , така што после активирањето на овој премин веќе логираниот корисник пристапил кон прегледување на својата библиотека прегледување на страна од категорија  $C$ , односно жетонот е преместен од местото  $C$  во местото  $ML$ .
- Кога корисникот се одлучил за да ја заврши својата посета после последен преглед на својата кошничка, се активира преминот  $t_{E\_ML}$  со интензитет  $\mu$ , така што жетонот е преместен од местото  $C$  во местото  $E$ .
- Кога се активира преминот  $t_{C\_cont}$ , интензитетот е  $\theta$ , така што после активирањето на овој премин корисникот повторно пристапил кон страна од категорија  $C$ .

Кога жетонот е во местото  $E$  нема овозможени премини, односно се смета дека корисникот ја завршил својата посета.

#### 4.2.2 Анализа на предложениот модел

Развиената мрежа е безбедна или врзана бидејќи секое нејзино место е врзано или безбедно, соодветно. Но, и во овој модел постои и т.н. deadlock ситуација, така што најкратките патеки до deadlock се:  $tA\_cont \rightarrow tE\_A$ ,  $tB \rightarrow tE\_B$ ,  $tL \rightarrow tE\_B$ ,  $tML \rightarrow tE\_ML$ ,  $tC \rightarrow tE\_C$ ,  $tD1 \rightarrow tE\_D1$ ,  $tD2 \rightarrow tE\_D2$ ,  $tD3 \rightarrow tE\_D3$  и  $tD4 \rightarrow tE\_D4$ .

Исто така се наидува и на конфликтна ситуација, и тоа во момент кога жетонот е во местата:  $A$ ,  $B$ ,  $L$ ,  $ML$ ,  $C$ ,  $D1$ ,  $D2$ ,  $D3$  и  $D4$ . Како што е објаснато во претходната секција, во секоја од овие ситуации, овозможени се повеќе од еден премин, така што системот се доведува до состојба на конфликт. Со активирање на еден од овие премини, останатите не можат да се активираат, жетонот е преместен во друго соодветно место и повторно системот се доведува до состојба на конфликт, која единствено се решава со активирање на една од премините:  $tE\_A$ ,  $tE\_B$ ,  $tE\_L$ ,  $tE\_ML$ ,  $tE\_C$ ,  $tE\_D1$ ,  $tE\_D2$ ,  $tE\_D3$ ,  $tE\_D4$ .

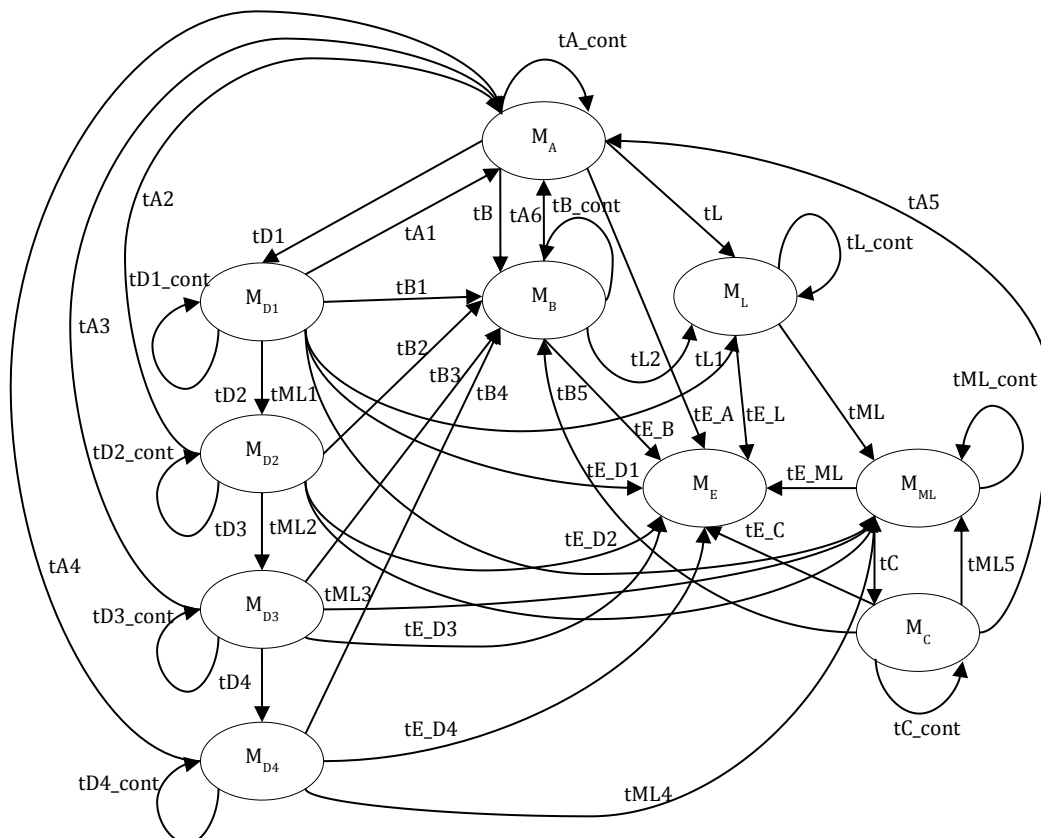
Моделот се воведува во состојба на секвенцијално извршување во повеќе случаи, како на пример:

- $tC$  може да се активира единствено само после активирање на премините:  $tML$ ,  $tML1$ ,  $tML2$ ,  $tML3$ ,  $tML4$ ,  $tML\_cont$ , или  $tML5$ .
- $tD2$  може да се активира единствено само после активирање на премините:  $tD1$  или  $tD1\_cont$ .
- $tD3$  може да се активира единствено само после активирање на премините:  $tD2$  или  $tD2\_cont$ .
- $tD4$  може да се активира единствено само после активирање на премините:  $tD3$  или  $tD3\_cont$ .
- итн.

И во овој случај водејќи се според [1], [45], врз развиениот модел реализирана е анализа на достижност со што се откриваат и земаат во предвид сите достапни означувања. Иградениот граф на достижност е прикажан на Слика 9. Иницијалното означување е  $M_A$ , а останатите достапни означувања се прикажани во Табела 13. И во овој случај, не постојат исчезнувачки означувања. Сите означувања се допирливи, што го претвора моделот во SPN мрежа, но секоја таква мрежа е GSPN мрежа.

**Табела 13 Означувања во GSPN мрежата**

Marking	A	B	E	L	ML	C	D1	D2	D3	D4
$M_A$	1	0	0	0	0	0	0	0	0	0
$M_B$	0	1	0	0	0	0	0	0	0	0
$M_E$	0	0	1	0	0	0	0	0	0	0
$M_L$	0	0	0	1	0	0	0	0	0	0
$M_{ML}$	0	0	0	0	1	0	0	0	0	0
$M_C$	0	0	0	0	0	1	0	0	0	0
$M_{D1}$	0	0	0	0	0	0	1	0	0	0
$M_{D2}$	0	0	0	0	0	0	0	1	0	0
$M_{D3}$	0	0	0	0	0	0	0	0	1	0
$M_{D4}$	0	0	0	0	0	0	0	0	0	1





И во овој случај, за опишување на влезните и излезните функции на моделот се користат матрици. Според секција 2.3, генерираните матрици се прикажани во Табела 14, Табела 15 и Табела 16, соодветно. Димензијата на сите овие матрици е  $44 \times 10$ , што е еквивалентно на  $n_t \times n_p$ , каде  $n_t$  е бројот на премини, а  $n_p$  е бројот на места.

**Табела 14 Forward incidence матрица I<sup>+</sup>**

	A	B	C	D1	D2	D3	D4	E	L	ML
tA_cont	1	0	0	0	0	0	0	0	0	0
tA2	1	0	0	0	0	0	0	0	0	0
tA3	1	0	0	0	0	0	0	0	0	0
tA4	1	0	0	0	0	0	0	0	0	0
tA5	1	0	0	0	0	0	0	0	0	0
tA6	1	0	0	0	0	0	0	0	0	0
tB	0	1	0	0	0	0	0	0	0	0
tB_cont	0	1	0	0	0	0	0	0	0	0
tB1	0	1	0	0	0	0	0	0	0	0
tB2	0	1	0	0	0	0	0	0	0	0
tB3	0	1	0	0	0	0	0	0	0	0
tB4	0	1	0	0	0	0	0	0	0	0
tB6	0	1	0	0	0	0	0	0	0	0
tC	0	0	1	0	0	0	0	0	0	0
tC_cont	0	0	1	0	0	0	0	0	0	0
tD1	0	0	0	1	0	0	0	0	0	0
tD1_cont	0	0	0	1	0	0	0	0	0	0
tD2	0	0	0	0	1	0	0	0	0	0
tD2_cont	0	0	0	0	1	0	0	0	0	0
tD3	0	0	0	0	0	1	0	0	0	0
tD3_cont	0	0	0	0	0	1	0	0	0	0
tD4	0	0	0	0	0	0	1	0	0	0
tD4_cont	0	0	0	0	0	0	1	0	0	0
tE_A	0	0	0	0	0	0	0	1	0	0
tE_B	0	0	0	0	0	0	0	1	0	0
tE_C	0	0	0	0	0	0	0	1	0	0
tE_D1	0	0	0	0	0	0	0	1	0	0
tE_D2	0	0	0	0	0	0	0	1	0	0
tE_D3	0	0	0	0	0	0	0	1	0	0
tE_D4	0	0	0	0	0	0	0	1	0	0
tE_L	0	0	0	0	0	0	0	1	0	0
tE_ML	0	0	0	0	0	0	0	1	0	0
tL	0	0	0	0	0	0	0	0	1	0
tL_cont	0	0	0	0	0	0	0	0	1	0
tL1	0	0	0	0	0	0	0	0	1	0
tL2	0	0	0	0	0	0	0	0	1	0
tML_cont	0	0	0	0	0	0	0	0	0	1
tML	0	0	0	0	0	0	0	0	0	1

tML1	0	0	0	0	0	0	0	0	0	1
tML2	0	0	0	0	0	0	0	0	0	1
tML3	0	0	0	0	0	0	0	0	0	1
tML4	0	0	0	0	0	0	0	0	0	1
tML6	0	0	0	0	0	0	0	0	0	1
tA1	1	0	0	0	0	0	0	0	0	0

**Табела 15 Backward incidence матрица I'**

	A	B	C	D1	D2	D3	D4	E	L	ML
tA_cont	1	0	0	0	0	0	0	0	0	0
tA2	0	0	0	0	1	0	0	0	0	0
tA3	0	0	0	0	0	1	0	0	0	0
tA4	0	0	0	0	0	0	1	0	0	0
tA5	0	0	1	0	0	0	0	0	0	0
tA6	0	1	0	0	0	0	0	0	0	0
tB	1	0	0	0	0	0	0	0	0	0
tB_cont	0	1	0	0	0	0	0	0	0	0
tB1	0	0	0	1	0	0	0	0	0	0
tB2	0	0	0	0	1	0	0	0	0	0
tB3	0	0	0	0	0	1	0	0	0	0
tB4	0	0	0	0	0	0	1	0	0	0
tB6	0	0	1	0	0	0	0	0	0	0
tC	0	0	0	0	0	0	0	0	0	1
tC_cont	0	0	1	0	0	0	0	0	0	0
tD1	1	0	0	0	0	0	0	0	0	0
tD1_cont	0	0	0	1	0	0	0	0	0	0
tD2	0	0	0	1	0	0	0	0	0	0
tD2_cont	0	0	0	0	1	0	0	0	0	0
tD3	0	0	0	0	1	0	0	0	0	0
tD3_cont	0	0	0	0	0	1	0	0	0	0
tD4	0	0	0	0	0	1	0	0	0	0
tD4_cont	0	0	0	0	0	0	1	0	0	0
tE_A	1	0	0	0	0	0	0	0	0	0
tE_B	0	1	0	0	0	0	0	0	0	0
tE_C	0	0	1	0	0	0	0	0	0	0
tE_D1	0	0	0	1	0	0	0	0	0	0
tE_D2	0	0	0	0	1	0	0	0	0	0
tE_D3	0	0	0	0	0	1	0	0	0	0
tE_D4	0	0	0	0	0	0	1	0	0	0
tE_L	0	0	0	0	0	0	0	0	1	0
tE_ML	0	0	0	0	0	0	0	0	0	1
tL	1	0	0	0	0	0	0	0	0	0
tL_cont	0	0	0	0	0	0	0	0	1	0
tL1	0	0	0	1	0	0	0	0	0	0
tL2	0	1	0	0	0	0	0	0	0	0
tML_cont	0	0	0	0	0	0	0	0	0	1

tML	0	0	0	0	0	0	0	0	1	0
tML1	0	0	0	1	0	0	0	0	0	0
tML2	0	0	0	0	1	0	0	0	0	0
tML3	0	0	0	0	0	1	0	0	0	0
tML4	0	0	0	0	0	0	1	0	0	0
tML6	0	0	1	0	0	0	0	0	0	0
tA1	0	0	0	1	0	0	0	0	0	0

**Табела 16 Комбинирана incidence матрица I**

	A	B	C	D1	D2	D3	D4	E	L	ML
tA_cont	0	0	0	0	0	0	0	0	0	0
tA2	1	0	0	0	-1	0	0	0	0	0
tA3	1	0	0	0	0	-1	0	0	0	0
tA4	1	0	0	0	0	0	-1	0	0	0
tA5	1	0	-1	0	0	0	0	0	0	0
tA6	1	-1	0	0	0	0	0	0	0	0
tB	-1	1	0	0	0	0	0	0	0	0
tB_cont	0	0	0	0	0	0	0	0	0	0
tB1	0	1	0	-1	0	0	0	0	0	0
tB2	0	1	0	0	-1	0	0	0	0	0
tB3	0	1	0	0	0	-1	0	0	0	0
tB4	0	1	0	0	0	0	-1	0	0	0
tB6	0	1	-1	0	0	0	0	0	0	0
tC	0	0	1	0	0	0	0	0	0	-1
tC_cont	0	0	0	0	0	0	0	0	0	0
tD1	-1	0	0	1	0	0	0	0	0	0
tD1_cont	0	0	0	0	0	0	0	0	0	0
tD2	0	0	0	-1	1	0	0	0	0	0
tD2_cont	0	0	0	0	0	0	0	0	0	0
tD3	0	0	0	0	-1	1	0	0	0	0
tD3_cont	0	0	0	0	0	0	0	0	0	0
tD4	0	0	0	0	0	-1	1	0	0	0
tD4_cont	0	0	0	0	0	0	0	0	0	0
tE_A	-1	0	0	0	0	0	0	1	0	0
tE_B	0	-1	0	0	0	0	0	1	0	0
tE_C	0	0	-1	0	0	0	0	1	0	0
tE_D1	0	0	0	-1	0	0	0	1	0	0
tE_D2	0	0	0	0	-1	0	0	1	0	0
tE_D3	0	0	0	0	0	-1	0	1	0	0
tE_D4	0	0	0	0	0	0	-1	1	0	0
tE_L	0	0	0	0	0	0	0	1	-1	0
tE_ML	0	0	0	0	0	0	0	1	0	-1
tL	-1	0	0	0	0	0	0	0	1	0
tL_cont	0	0	0	0	0	0	0	0	0	0
tL1	0	0	0	-1	0	0	0	0	1	0
tL2	0	-1	0	0	0	0	0	0	1	0



## 2. Конструирање на infinitesimal generator матрица

Соодветната infinitesimal generator матрица  $Q$  конструирана со нумерирање на состојбите прикажани на Табела 13, за соодветниот СТМС е прикажана подолу:

$$Q = \begin{bmatrix} x_1 & \lambda & \mu & \kappa & 0 & 0 & \varepsilon & 0 & 0 & 0 \\ \alpha & x_2 & \mu & \kappa & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mu & x_3 & \nu & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mu & 0 & x_4 & \theta & 0 & 0 & 0 & 0 \\ \alpha & \lambda & \mu & 0 & \nu & x_5 & 0 & 0 & 0 & 0 \\ \alpha & \lambda & \mu & \kappa & \nu & 0 & x_6 & \gamma & 0 & 0 \\ \alpha & \lambda & \mu & 0 & \nu & 0 & 0 & x_7 & \delta & 0 \\ \alpha & \lambda & \mu & 0 & \nu & 0 & 0 & 0 & x_8 & \beta \\ \alpha & \lambda & \mu & 0 & \nu & 0 & 0 & 0 & 0 & x_9 \end{bmatrix} \quad (4.68)$$

каде,

$$\begin{aligned} x_1 &= -(\lambda + \mu + \kappa + \varepsilon + \alpha - \alpha) \\ x_2 &= -(\alpha + \mu + \kappa + \lambda - \lambda) \\ x_3 &= -(\mu + \nu + \kappa - \kappa) \\ x_4 &= -(\mu + \theta + \nu - \nu) \\ x_5 &= -(\alpha + \lambda + \mu + \nu + \theta - \theta) \\ x_6 &= -(\alpha + \lambda + \mu + \kappa + \nu + \gamma + \varepsilon - \varepsilon) \\ x_7 &= -(\alpha + \lambda + \mu + \nu + \delta + \gamma - \gamma) \\ x_8 &= -(\alpha + \lambda + \mu + \nu + \beta + \delta - \delta) \\ x_9 &= -(\alpha + \lambda + \mu + \nu + \beta - \beta) \end{aligned} \quad (4.69)$$

## 3. Решавање на веројатносниот вектор на стационарна состојба

Во соодветниот СТМС, има десет состојби за кои елементите на векторот  $\vec{\pi}$  се добиваат преку решавање на следниот систем линеарни равенки:

$$\left\{ \begin{array}{l} [\pi_0, \pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \pi_6, \pi_7, \pi_8, \pi_9] \begin{bmatrix} x_1 & \lambda & \mu & \kappa & 0 & 0 & \varepsilon & 0 & 0 & 0 \\ \alpha & x_2 & \mu & \kappa & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mu & x_3 & \nu & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mu & 0 & x_4 & \theta & 0 & 0 & 0 & 0 \\ \alpha & \lambda & \mu & 0 & \nu & x_5 & 0 & 0 & 0 & 0 \\ \alpha & \lambda & \mu & \kappa & \nu & 0 & x_6 & \gamma & 0 & 0 \\ \alpha & \lambda & \mu & 0 & \nu & 0 & 0 & x_7 & \delta & 0 \\ \alpha & \lambda & \mu & 0 & \nu & 0 & 0 & 0 & x_8 & \beta \\ \alpha & \lambda & \mu & 0 & \nu & 0 & 0 & 0 & 0 & x_9 \end{bmatrix} = 0 \\ \pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6 + \pi_7 + \pi_8 + \pi_9 = 1 \end{array} \right. \quad (4.70)$$

Решавање на веројатносниот вектор на стационарна состојба и останатите чекори од алгоритмот кои се однесуваат на пресметка на соодветни мерки за мерење на перформанси, се реализирани во следната секција, врз студија на случај.

#### 4.2.4 Евалуација на мерки на перформанси преку студија на случај

Навигациските податоци се собирани на серверска страна во одредена временска рамка и се во стандарден W3C<sup>9</sup> формат, а се однесуваат на историски податоци за прегледани страни во рамки на електронската книжарница. Изворната датотека содржи околу 415000 записи, така што секој запис содржи податоци кои припаѓаат на повеќе категории (атрибути), но за ова истражување, од посебно значење се категориите: корисник, датум и време и посетена страна.

Пред да се премине на потребните анализи, потребно беше да се направи пред-процесирање и чистење на податоците, со цел нивно претворање во формат лесен за интерпретирање и понатамошна обработка. Ова подразбира отстранување на сите некомплетни записи, односно записи на кои им недостасува некој од клучните атрибути или истите се неконзистентни. Вака исчистените податоци, претставуваат влез во скрипта, со чија помош динамички се идентификуваат корисниците, сите кориснички посети, нивното времетраење и се утврдуваат последователните прегледи на страни, по корисник во рамките на секоја посета. Исто така, со помош на скриптата, се прави статистичка анализа на податоците, со цел откривање на шемите на општото однесување на корисниците. После чекорот на пред-процесирање и чистење на податоците, вкупниот број на кориснички прегледи на страни е околу 322000. Исто така, идентификувани се 1984 различни (уникатни) корисници и 15433 различни посети. Овие податоци, како и податоците кои се однесуваат на: време помеѓу две последователни посети, бројот на посети по корисник и бројот на прегледи по посета, илустрирани се во Табела 17.

Табела 17 Преглед на податоците од статистичката анализа

Статистичка анализа на лог податоците				
Вкупен број на прегледи	321854			
Вкупен број на корисници	1984			
Вкупен број на посети	15433			
	<b>Mean</b>		<b>(Min, Max)</b>	<b>SD</b>
Време помеѓу две последователни посети	3d:20h:06m:59s	(0d:0h:0m:0s, 332d:23h:52m:57s)		
Број на посети по корисник	7.78		(1, 229)	13.89
Број на прегледи по посета	20.85		(1, 2929)	52.60

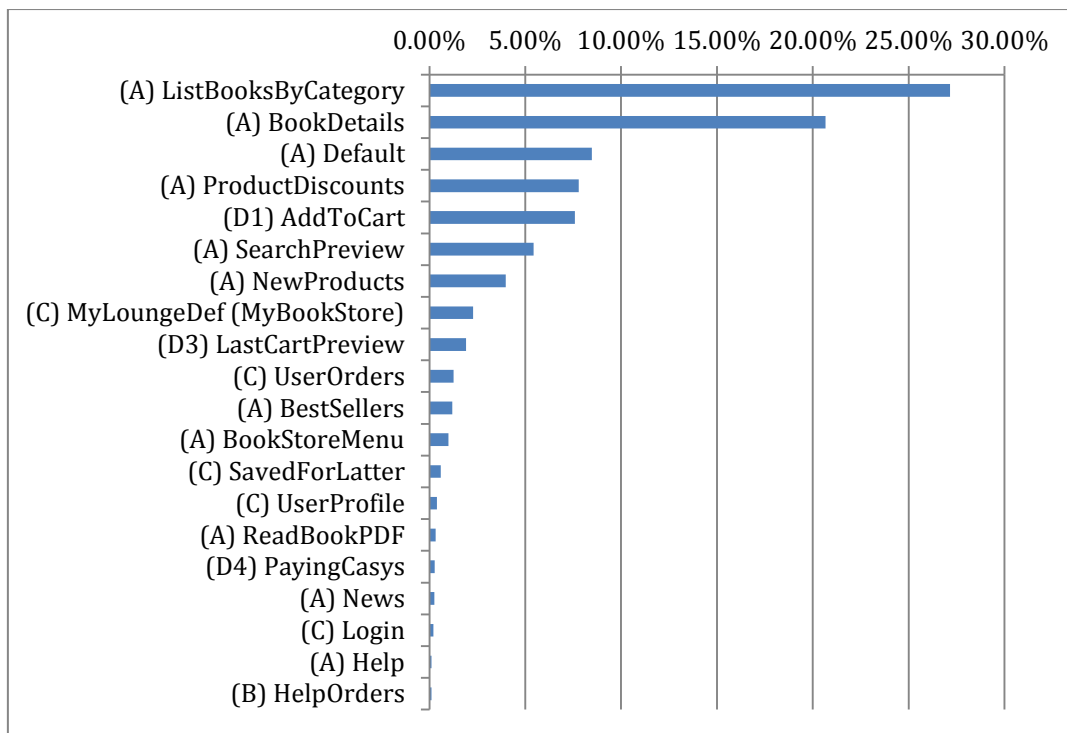
<sup>9</sup> <http://www.w3.org/Daemon/User/Config/Logging.html#common-logfile-format>

Врз основа на структурата на електронската книжарница и посетите регистрирани во лог датотеката, во Табела 18 прикажан е бројот на најчесто прегледувани страни (линкови).

**Табела 18 Најчесто прегледувани страни (линкови) во лог датотеката**

<b>Веб страна</b>	<b>%</b>
(A) ListBooksByCategory	27.17%
(A) BookDetails	20.66%
(A) Default	8.47%
(A) ProductDiscounts	7.80%
(D1) AddToCart	7.59%
(A) SearchPreview	5.44%
(A) NewProducts	3.98%
(C) MyLoungeDef (MyBookStore)	2.27%
(D3) LastCartPreview	1.91%
(C) UserOrders	1.26%
(A) BestSellers	1.19%
(A) BookStoreMenu	0.99%
(C) SavedForLatter	0.58%
(C) UserProfile	0.39%
(A) ReadBookPDF	0.32%
(D4) PayingCasys	0.28%
(A) News	0.26%
(C) Login	0.20%
(A) Help	0.12%
(B) HelpOrders	0.11%

Од визуелната инспекција на Слика 11, може да се утврди дека најголемиот број кориснички навигациски акции во рамки на лог датотеката се однесуваат на преглед на дадена книга по категорија (27.17%), преглед на детали за избрана книга (20.66%), како и прегледување на почетната страна (8.47%). Приближно еднаков број на прегледи припаѓаат на прегледот на попусти (7.80%) и ставањето на книги во кошничка (7.59%). Дobar дел од корисничките акции претставуваат пребарување на дадена книга по различни критериуми (5.44%) и преглед на нови производи (3.98%). Исто така може да се забележи дека мал дел (1.91%) од прегледите се однесуваат на последен преглед на кошничката, но уште помал дел (0.28%) претставуваат купување на дадена книга, односно плаќање, како последен чекор во процесот на електонско купување.



**Слика 11** Визуелен приказ на бројот на најчесто посетувани страни (линкови)

Исто така, врз основа на структурата на електронската книжарница и корисничките акции на уникатните корисници, регистрирани во лог датотеката, во Табела 19 прикажан е бројот на посети на најчесто посетуваните страни од страна на уникатните корисници.

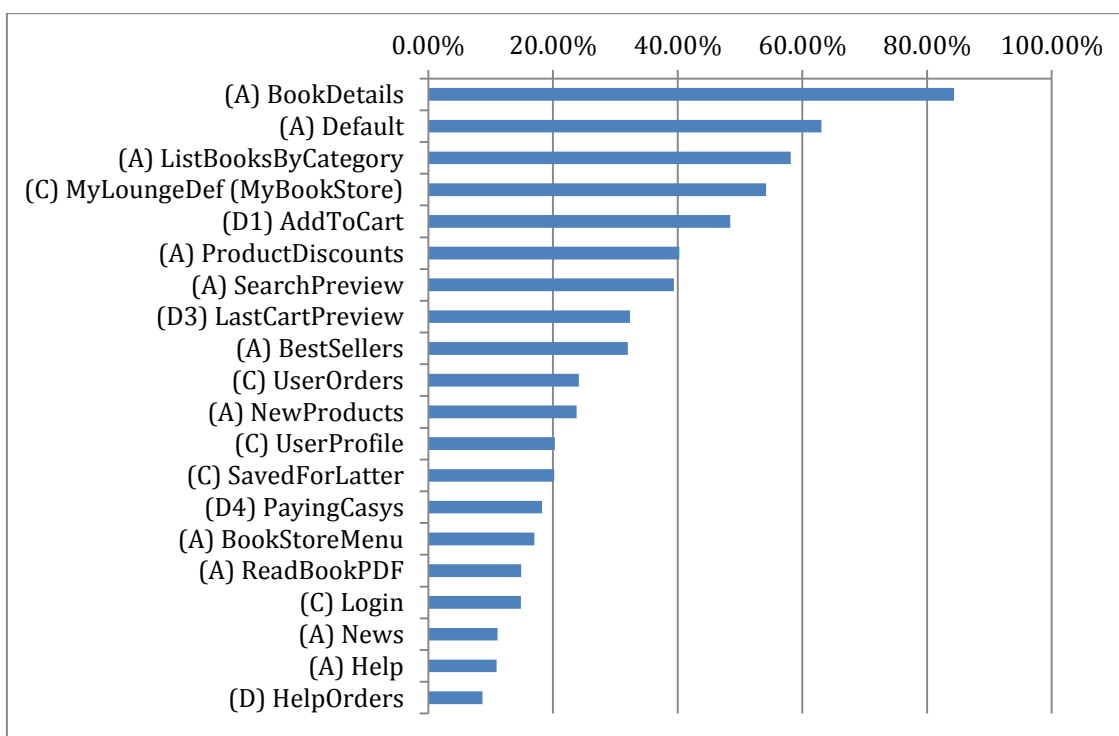
**Табела 19** Најчесто посетувани страни (линкови) од уникатни корисници

Линкови (Веб страни)	%
(A) BookDetails	84.32%
(A) Default	63.05%
(A) ListBooksByCategory	58.17%
(C) MyLoungeDef (MyBookStore)	54.18%
(D1) AddToCart	48.44%
(A) ProductDiscounts	40.27%
(A) SearchPreview	39.42%
(D3) LastCartPreview	32.36%
(A) BestSellers	32.01%
(C) UserOrders	24.14%
(A) NewProducts	23.79%
(C) UserProfile	20.31%
(C) SavedForLatter	20.21%
(D4) PayingCasys	18.25%
(A) BookStoreMenu	17.04%
(A) ReadBookPDF	14.92%
(C) Login	14.87%



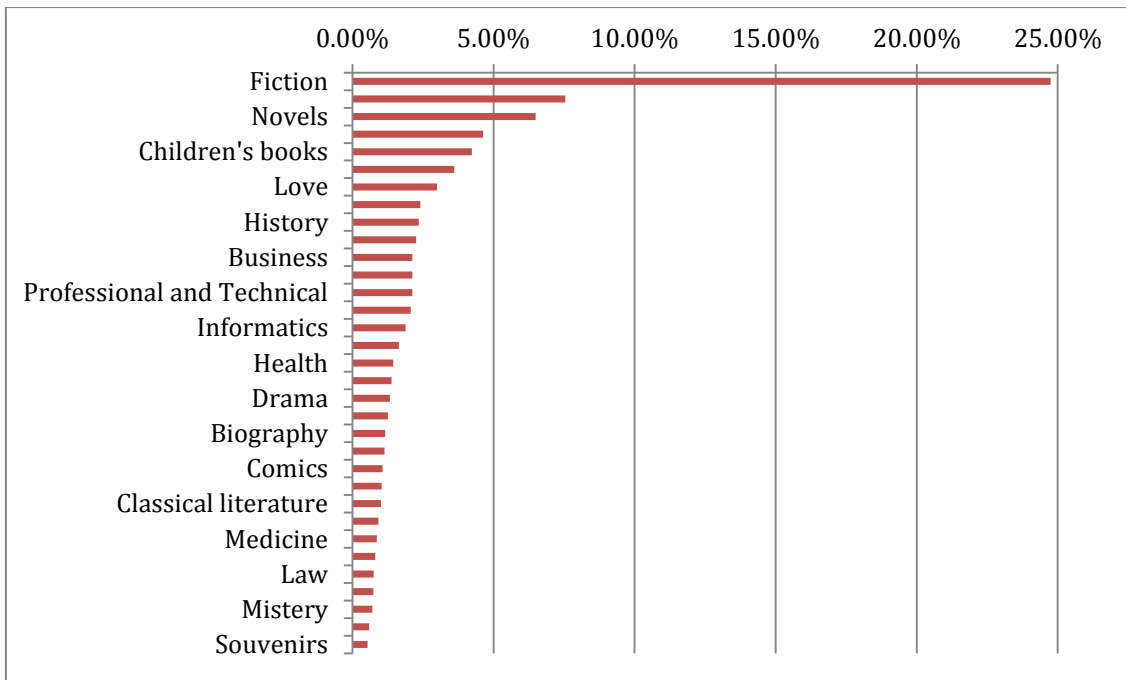
(A) News	11.14%
(A) Help	10.94%
<u>(D) HelpOrders</u>	<u>8.72%</u>

Од визуелната инспекција на Слика 12, може да се утврди дека најголем број од корисниците (84.32%), барем еднаш во рамки на нивните посети, се заинтересирани за преглед на детали на избрана книга или (63.05%) ја прегледале почетната страна на електронската книжарница. Исто така, голем број корисници (58.17%) заинтересирани се за филтрирање на книги по категорија или за прегледување на својата библиотека (54.18%), од каде имаат пристап до своите лични податоци, своите нарачки и сл. Уште еднаш, мал дел од корисниците (18.25%) барем еднаш успешно ја завршиле својата нарачка (купиле книга), наспроти скоро половина од корисниците (48.44%) кои додале книга во својата кошничка.



**Слика 12 Визуелен приказ на бројот на најчесто посетувани линкови од уникатни корисници**

Бидејќи најголем дел од вкупниот број на посети (27.17%) припаѓа на корисничката акција – филтрирање на книги (произоди) по категорија и исто така, голем број корисници (58.17%) се заинтересирани за оваа опција, направена е анализа на најпопуларни категории и истата е прикажана на Слика 13.



Слика 13 Анализа на најпопуларни категории на книги (производи)

#### 4.2.4.1 Кластерирање на секвенци – концепт и примена

За кластерирање на однесувањето на корисниците при процесот на навигација, го применет е алгоритмот за кластерирање секвенци, детално прикажан во [40]. Овој алгоритам преку комбинација од стандардни методи за кластерирање и техники за анализа на секвенци, засновани на Марков Ланец, групира корисничко однесување според сличноста на секвенцијалниот редослед на кориснички акции. Одлуката за овој пристап, е бидејќи корисна информација за однесувањето на корисниците освен самите премини носи и нивниот редослед. Под секвенца се подразбира серија на дискретни настани, кои во овој случај претставуваат кориснички прегледи на веб страни во рамки на електронската книжарница.

Прв чекор на алгоритмот е креирање на модел, во кој како влез се задаваат следните две табели:

- главна табела (case table), која во суштина претставува стандардна табела каде се чуваат податоци за уникатните корисниците, и
- вгнездена табела (nested table), како специјален тип на табела во која ги чуваме секвенците кориснички акции, најчесто претставени преку: идентификатор на корисници, идентификатор на секвенци кој всушност се однесува на секвенцијалниот редослед и претставува клуч на оваа табела и корисничка акција.

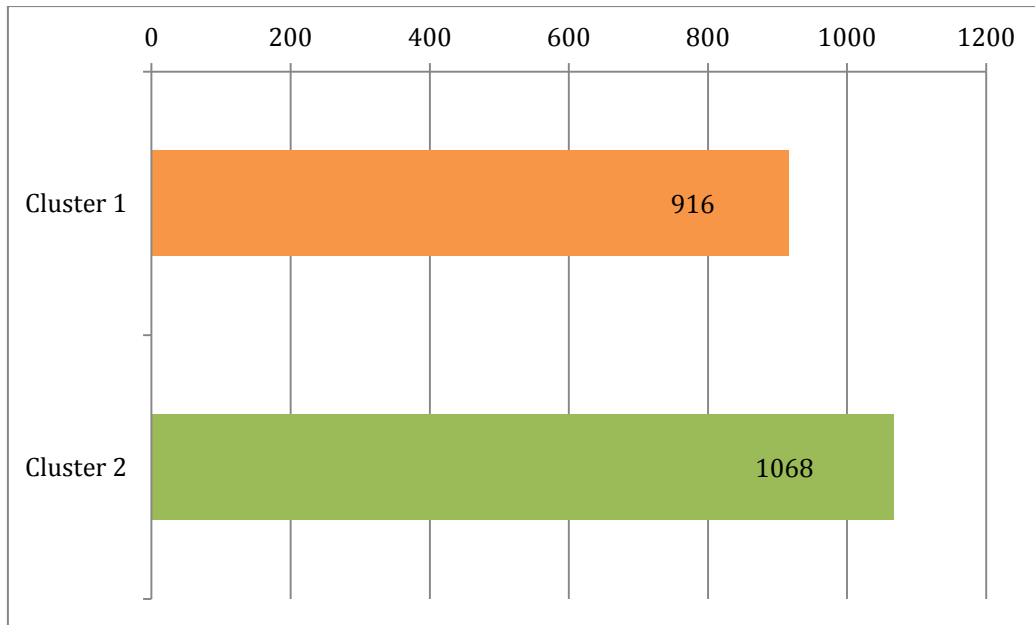
Втор чекор на алгоритмот е иницијализирање на соодветните параметри за кластерирање, со чија помош се определуваат и контролираат влезните податоци и повратните резултати на алгоритмот:

- CLUSTER\_COUNT – се однесува на бројот на кластери што ќе ги генерира алгоритмот. Во овој случај беше избрана опцијата автоматско предвидување на оптималниот број кластери според влезното множество податоци (CLUSTER\_COUNT=0).
- MINIMUM\_SUPPORT – се однесува на минималниот број на записи во секој кластер, со цел да се избегне непотребно формирање на многу кластери. Овој параметар во овде се користи со својата default вредност (MINIMUM\_SUPPORT=10).
- MAXIMUM\_STATES – се однесува на максималниот број на состојби во рамки на атрибутот според кој се прави кластерирањето. Во овој случај беше избрана default опцијата (MAXIMUM\_STATES=100). Доколку влезните податоци содржат атрибут кој се кластерира со повеќе од зададениот бројот на состојби, автоматски се прави feature selection според популарноста на состојбите.
- MAXIMUM\_SEQUENCE\_STATES – се однесува на максималниот број на состојби во рамки на секвенцијалниот атрибут, не повеќе од 100, бидејќи со зголемување на бројот на состојби ( $M$ ) пропорционално се зголемува потребното време за процесирање на податоците ( $M^2$ ). И во овој случај, доколку влезните податоци содржат секвенцијални атрибути со повеќе од зададениот бројот на состојби, автоматски се прави feature selection според популарноста на состојбите. Беше одлучено зададениот бројот на состојби да е 100 (MAXIMUM\_SEQUENCE\_STATES=100).

Следно, се процесираат податоците според зададените параметри, се формираат кластери, така што секој запис со одредена веројатност припаѓа на секој од кластерите и се ре-евалуира моделот, се додека истиот не конвергира.

#### 4.2.4.2 Интерпретирање на кластерите

Според поставените влезни параметри на алгоритмот и влезното множество податоци, однесувањето на корисниците при процесот на навигација, автоматски е групирано во два кластери, како најоптимално решение на овој проблем. Како што е илустрирано на Слика 14, бројот на корисници во секој од кластерите е приближно еднаков.



Слика 14 Број на предвидени корисници по кластер

Визуелен приказ на профилите на кластерите е даден во Табела 20, така што секоја колона претставува кластер, а секој ред претставува секвенцијален атрибут. Во секоја ќелија прикажан е хистограм на секвенци, кои се однесуваат на секвенцијалните кориснички акции.

Табела 20 Профили на кластери

Атрибут	Популација	Кластер 1	Кластер 2
List BooksforCategory	[Хистограм]	[Хистограм]	[Хистограм]
BookDetails	[Хистограм]	[Хистограм]	[Хистограм]
Default	[Хистограм]	[Хистограм]	[Хистограм]
ProductDiscounts	[Хистограм]	[Хистограм]	[Хистограм]
SearchPreview	[Хистограм]	[Хистограм]	[Хистограм]
AddToCart	[Хистограм]	[Хистограм]	[Хистограм]
NewProducts	[Хистограм]	[Хистограм]	[Хистограм]
Publisher	[Хистограм]	[Хистограм]	[Хистограм]
Other	[Хистограм]	[Хистограм]	[Хистограм]

На прв поглед, во Табела 20 може да се забележи дека во првиот кластер преовладуваат секвенци во кои доминира прегледување на страната што се однесува на детали за дадена книга, додека пак во вториот кластер преовладуваат секвенци во кои доминира прегледување на страната што се однесува на преглед на книги по дадена категорија. Сепак, детален приказ на карактеристите на двата кластери е даден на Слика 15 и Слика

16, соодветно. За секој атрибут, односно вредност на тој атрибут, прикажана е фреквенцијата, односно веројатноста со која тој е застапен во дадениот кластер. Поточно, илустрирани се дел од најчестите состојби, односно секвенци на кориснички акции, карактеристични за тој кластер, сортирани по опаѓачки редолед.

Variables	Values	Probability
URL Category.Transitions	[Start] -> BookDetails	
URL Category	BookDetails	
URL Category.Transitions	[Start] -> Default	
URL Category.Transitions	[Start] -> MyLoungeDef	
URL Category	ListBooksforCategory	
URL Category.Transitions	ListBooksforCategory,ListBooksforC...	
URL Category.Transitions	BookDetails,BookDetails	
URL Category	Default	
URL Category	ProductDiscounts	
URL Category	SearchPreview	
URL Category	AddToCart	
URL Category.Transitions	ProductDiscounts,ProductDiscounts	
URL Category	NewProducts	
URL Category.Transitions	[Start] -> AddToCart	
URL Category	Publisher	
URL Category.Transitions	Default,Default	
URL Category.Transitions	Publisher,Publisher	
URL Category.Transitions	[Start] -> ListBooksforCategory	
URL Category.Transitions	SearchPreview,SearchPreview	
URL Category.Transitions	NewProducts,NewProducts	

Слика 15 Карактеристики на Кластер 1

Како што се гледа на Слика 15, најчеста вредност на секвенцијалниот атрибут е [Start]->BookDetails, што значи дека најголем дел од корисниците во првиот кластер ја започнуваат својата посета со преглед на детали за дадена книга. Од друга страна, се забележува дека добар дел од корисниците ја започнуваат својата посета отворајќи ја почетната страна на електронската книжарница – [Start]->Default или со прегледување на својата библиотека – [Start]->MyLoungeDef. Страната што се однесува на детали за дадена книга е најпопуларна меѓу овие корисници, по што следуваат страните: преглед на книги по дадена категорија, почетна, попусти, пребарување, ставање во кошничка, нови производи, издавачи, итн. Исто така, прикажани се премините од една на друга корисничка акција. Како најверојатен премин за повеќето корисници во овој кластер, може да се издвои преминот ListBooksforCategory -> ListBooksforCategory, што значи прегледување на страницата што се однесува на преглед на книги по дадена категорија, последователно. Следно, со помала веројатност би се случиле премините: прегледување на страната што се однесува на детали за дадена книга после прегледување на страницата што се однесува на преглед на книги, ListBooksforCategory -> BookDetails; последователни прегледи на страницата што се однесува на детали за дадена книга,

BookDetails -> BookDetails; последователни прегледи на страницата што се однесува на попусти, ProductDiscounts -> ProductDiscounts; итн.

Variables	Values	Probability
URL Category	ListBooksforCategory	
URL Category.Transitions	ListBooksforCategory,ListBooksforC...	
URL Category.Transitions	[Start] -> BookDetails	
URL Category.Transitions	[Start] -> MyLoungeDef	
URL Category	BookDetails	
URL Category.Transitions	[Start] -> AddToCart	
URL Category.Transitions	[Start] -> Default	
URL Category.Transitions	[Start] -> ListBooksforCategory	
URL Category	ProductDiscounts	
URL Category	AddToCart	
URL Category.Transitions	ListBooksforCategory,BookDetails	
URL Category.Transitions	ProductDiscounts,ProductDiscounts	
URL Category.Transitions	BookDetails,BookDetails	
URL Category.Transitions	BookDetails,ListBooksforCategory	
URL Category.Transitions	[Start] -> Login	
URL Category	Default	
URL Category	AddressEntry	
URL Category.Transitions	AddToCart,AddToCart	
URL Category	LastCartPreview	
URL Category	SearchPreview	

Слика 16 Карактеристики на Кластер 2

Од деталниот приказ на Слика 16, најчеста вредност на секвенцијалниот атрибут во вториот кластер е ListBooksforCategory, што значи дека најголем дел од корисниците во вториот кластер се заинтересирани за прегледување на страната што се однесува на преглед на книги по дадена категорија. Сортирани по популарноста, следуваат страните: детали за дадена книга, попусти, ставање во кошничка, почетна, промена на адреса за достава, последен преглед на кошничка, итн. Од друга страна, слично со првиот кластер, но со помала веројатност, се забележува дека добар дел од корисниците ја започнуваат својата посета со преглед на детали за дадена книга.

Исто така, посетите на помал дел од корисниците започнуваат со прегледување на својата библиотека – [Start]->MyLoungeDef или со ставање на книга во кошничка – [Start]-> AddToCart. Од прикажаното, како најверојатен премин за повеќето корисници во овој кластер, може да се издвои преминот ListBooksforCategory -> ListBooksforCategory, што значи прегледување на страницата што се однесува на преглед на книги по дадена категорија, последователно. Слично како во првиот кластер, следни премини со помала веројатност се: прегледување на страната што се однесува на детали за дадена книга после прегледување на страницата што се однесува на преглед на книги, ListBooksforCategory -> BookDetails; последователни прегледи на страницата што се

однесува на попусти, ProductDiscounts -> ProductDiscounts; последователни прегледи на страницата што се однесува на детали за дадена книга, BookDetails -> BookDetails; итн.

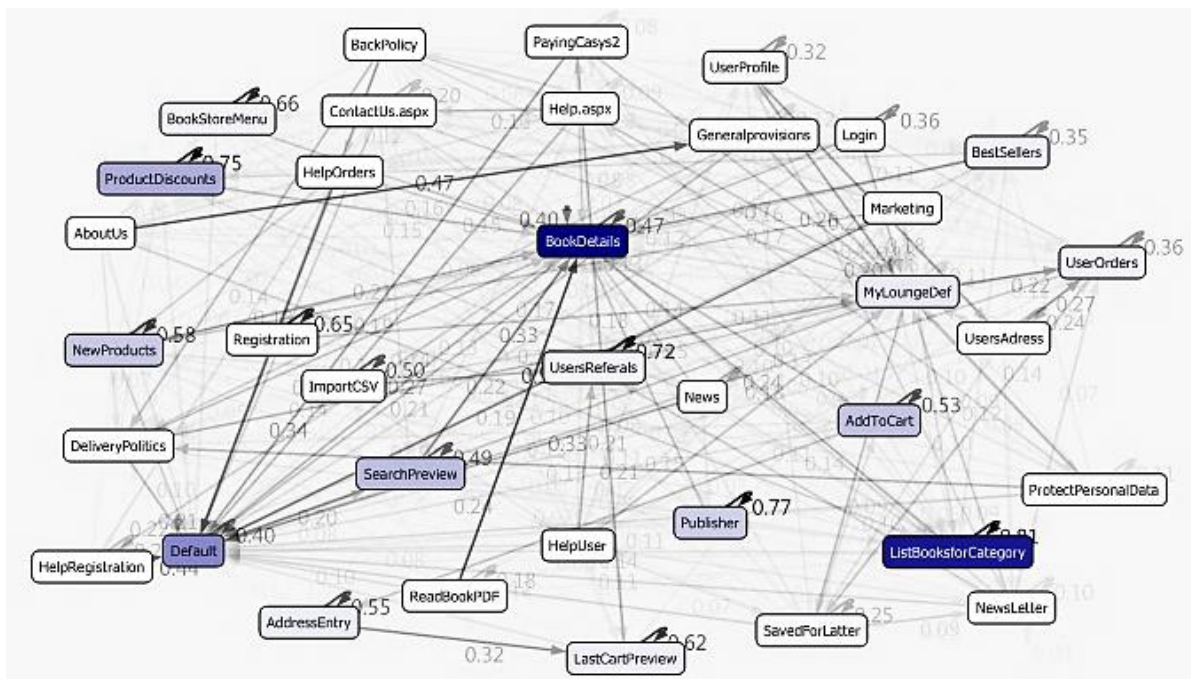
Variables	Values	Favors Cluster 1	Favors Cluster 2
URL Category.Transitions	AboutUs-> [End]		
URL Category.Transitions	Generalprovisions-> UserProfile		
URL Category.Transitions	ProtectPersonalData-> HelpRegistr...		
URL Category.Transitions	Marketing-> MyLoungeDef		
URL Category.Transitions	BackPolicy-> MyLoungeDef		
URL Category.Transitions	Publisher-> Publisher		
URL Category.Transitions	BookStoreMenu-> BookStoreMenu		
URL Category.Transitions	ImportCSV-> HelpUser		
URL Category.Transitions	ImportCSV-> ImportCSV		
URL Category.Transitions	BookStoreMenu-> ProductDiscounts		
URL Category.Transitions	AboutUs-> Generalprovisions		
URL Category.Transitions	HelpRegistration-> Default		
URL Category.Transitions	HelpRegistration-> BookDetails		
URL Category.Transitions	Marketing-> Default		
URL Category.Transitions	ImportCSV-> ProductDiscounts		
URL Category.Transitions	HelpUser-> [End]		
URL Category.Transitions	UsersReferrals-> UsersReferrals		
URL Category.Transitions	HelpRegistration-> ContactUs.aspx		
URL Category.Transitions	BookStoreMenu-> Help.aspx		
URL Category.Transitions	News-> News		

Слика 17 Дел од разликите меѓу двата кластери

На Слика 17 илустрирани се дел од разликите меѓу двата кластери. Една од најзначајните разлики е тоа што корисниците од вториот кластер ја завршуваат својата посета откако ја прегледале страната која содржи информации за електронската книжарница – AboutUs -> [End]. Исто така, корисниците од вториот кластер преферираат премин од страни од категорија В кон страни од категорија С, што се гледа во следниот случај: Generalprovisions -> UserProfile, Marketing -> MyLoungeDef, BackPolicy -> MyLoungeDef, и сл. Овие корисници почесто се одлучуваат да побараат помош за своите акции, што е забележано во следните ситуации: ProtectPersonalData -> HelpRegistration, ImportCSV -> HelpUser, BookStoreMenu -> Help, итн. Исто така, тие повеќе се интересираат за попусти, како што е прикажано со: BookStoreMenu -> ProductDiscounts, ImportCSV -> ProductDiscounts, итн. За разлика од нив, видно е дека корисниците од првиот кластер се интересираат за издавачи, книжарници и почетната страна, Publisher -> Publisher, BookStoreMenu -> BookStoreMenu, HelpRegistration -> Default, соодветно.

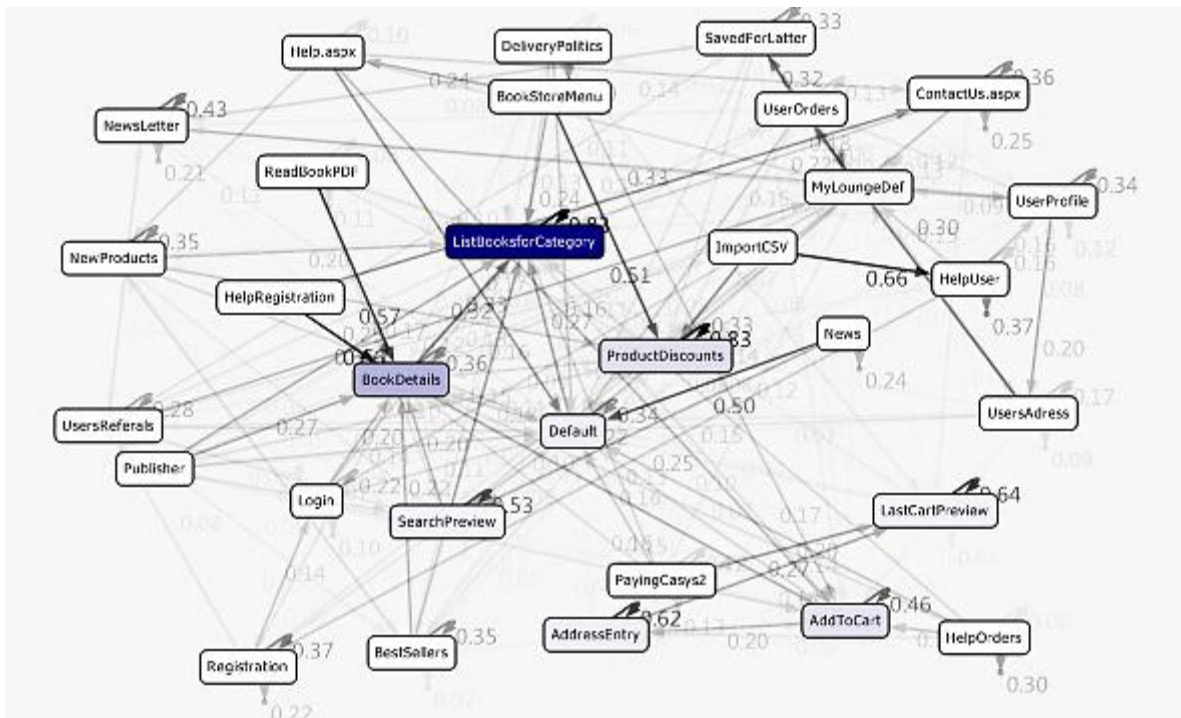
Детален приказ на премините и транзициските веројатности за двата кластери е даден на Слика 18 и Слика 19, соодветно за првиот и вториот кластер. Од овие дијаграми кои всушност претставуваат Маркови Ланци [40], лесно може да се утврдат шемите на однесување на корисниците при процесот на навигација. Секое теме претставува секвенцијална состојба, а секој раб претставува премин од една во друга состојба.

Работите се насочени и на секој раб му е придружена одредена тежина што претставува транзициска веројатност. Бојата на темето е поврзана со значајноста на состојбата што ја претставува, така што колку бојата е потемна, толку таа состојба е попопуларна меѓу корисниците во кластерот. Во таа насока, главен интерес на корисниците од првиот кластер се: преглед на детали за дадена книга, преглед на книги по дадена категорија, преглед на почетна страна, попусти, пребарување, нови книги, издавачи, ставање на книга во кошничка, книжарници, најпродавани книги, ставање во кошничка, избор на адреса за достава, последен преглед на кошничката, како и преглед на својата библиотека. Од друга страна, корисниците од вториот кластер се интересираат за: преглед на книги по дадена категорија, преглед на детали за дадена книга, попусти, ставање во кошничка, избор на адреса за достава, последен преглед на кошничката, како и преглед на својата библиотека.



Слика 18 Дијаграм на предни состојби во Кластер 1

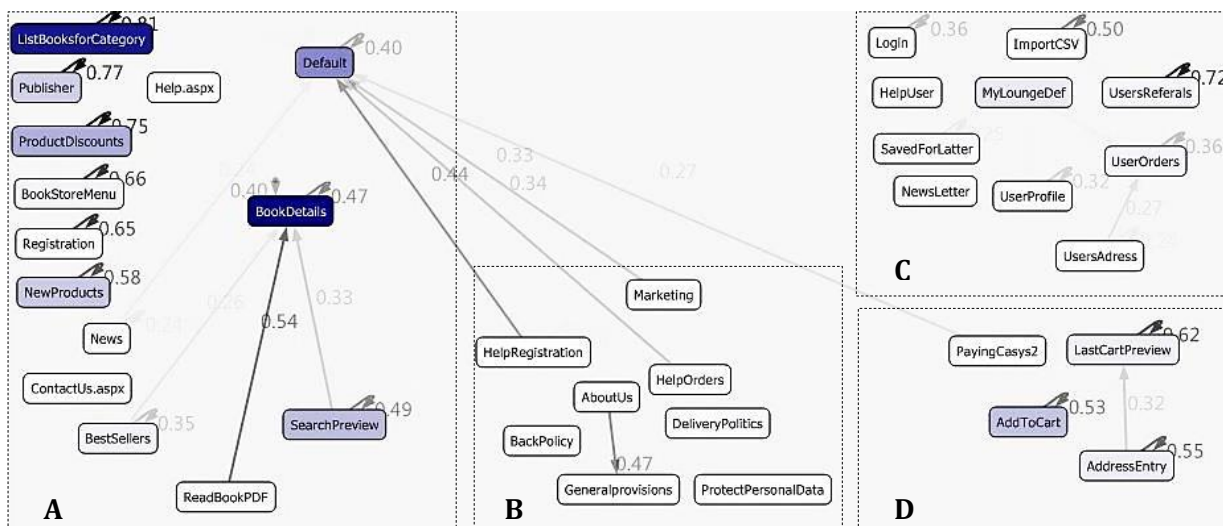




Слика 19 Дијаграм на предни состојби во Кластер 2

Поради комплексноста на транзициските дијаграми на Слика 18 и Слика 19, а со цел доловување на појасен приказ на шемите на однесување на корисниците при процесот на пребарување и решавање на моделите, прикажани се најкарактеристичните премини за првиот и вториот кластер, на Слика 20 и Слика 21 соодветно, групирани по категории на страни. Според илустрираната ситуација на Слика 20, издвоени се некои од откриените шеми на однесување на корисниците при процесот на навигација во првиот кластер:

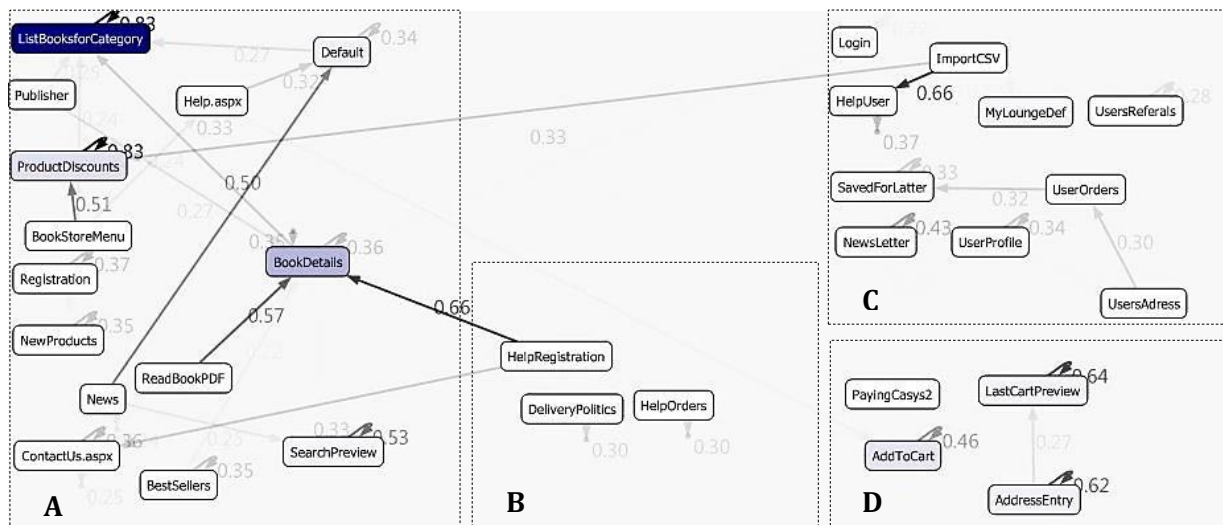
- 40% од корисниците ја започнуваат својата посета со преглед на детали за дадена книга. Скоро половина од тие корисниците (47%), се одлучуваат за повторна прегледување на истата.
- 22% од корисниците ја започнуваат својата посета со преглед на почетната страна, од кои, околу 40%, се одлучуваат за повторна прегледување на истата.
- Слична, околу 20% од корисниците ја започнуваат својата посета со прегледување на својата библиотека, а само 22% од нив ги прегледуваат своите нарачки.
- Една третина од корисниците што успешно купиле книга (27%), преминале на почетната страна.
- Слично, 32% од корисниците што промениле адреса на достава за последен пат ја погледнале својата кошничка. Но дури 62% од овие корисници, последователно ја погледнале својата кошничка.
- ИТН.



Слика 20 Најкарактеристични преодни состојби во Кластер 1

Исто така, аналогно на Слика 21, издвоени се некои од откриените шеми на однесување на корисниците при процесот на навигација во вториот кластер:

- 35% од корисниците ја започнуваат својата посета со преглед на детали за дадена книга. Една третина од тие корисниците (36%), се одлучуваат за повторна прегледување на истата.
- 22% од корисниците ја започнуваат својата посета со прегледување на својата библиотека.
- 66% од корисниците што поканиле пријатели, пристапиле кон страницата за помош на корисници. 37% од овие корисници, после оваа акција, ја завршиле својата посета.
- 27% од корисниците што промениле адреса на достава за последен пат ја погледнале својата кошничка. Но дури 64% од овие корисници, последователно ја погледнале својата кошничка, итн.



Слика 21 Најкарактеристични преодни состојби во Кластер 2

Вака претставените дијаграми на преодни состојби и нивните веројатности, не претставуваат Петриева мрежа. Тие едноставно даваат сликовита претстава за интересите на корисниците и нивното однесување во таа околина, што е од големо значење и помош за евалуирање на развиениот модел.

#### 4.2.4.3 Решавање на моделот

Според добиените квантитативни податоци од целокупните анализи, моделот е решен, валидиран и евалуиран во насока на утврдување на шемите на однесување на корисниците при навигација во рамки на првиот и вториот кластер. Поради комплексната структура на електронската книжарница, моделот, а со тоа и неговото решавање станува комплексно и сложено. Затоа, беше одлучено за негово решавање и илустрирање на примената на ново развиениот пристап, да бидат избрани дел од најпопуларните посетени страни по категории:

- Категорија А – BookDetails или Детали за даден производ (книга)
- Категорија В – DeliveryPolitics или Политика на испорака
- Категорија С – Login или Логирање, MyLounge или Моја Библиотека и UserOrders или Нарачки на корисникот
- Категорија D – AddToCart или Додади во кошничка, AddressEntry или Промена на адреса на достава, LastCartPreview или Последен предглед на кошничка и PayingCasus или Плаќање, како една целина во процесот на електронско купување.

##### 4.2.4.3.1 Резултати за Кластер 1

Пресметаните интензитети на сите временски премини во изградената GSPN мрежа, кои се однесуваат на првиот кластер, се прикажани во Табела 21.

**Табела 21 Интензитети на временските премини пресметани за Кластер 1**

	Rate
$\alpha$	0.000009
$\lambda$	0.000059
$\mu$	0.000003
$\kappa$	0.000004
$\nu$	0.000107
$\theta$	0.000882
$\varepsilon$	0.000073
$\gamma$	0.083333
$\beta$	0.047619

Веројатностите за активирање на временските премини се прикажани во Табела 22:

**Табела 22 Веројатности на активација на премините во Кластер 1**

<b>Probabilities</b>			
$P(tA\_cont)$	0.058346	$P(tD3)$	0.363144
$P(tB)$	0.401981	$P(tA3)$	0.000081
$P(tE\_A)$	0.020495	$P(tB3)$	0.000555
$P(tL)$	0.024440	$P(tE\_D3)$	0.000028
$P(tD1)$	0.494737	$P(tML3)$	0.001001
$P(tA6)$	0.115476	$P(tD3\_cont)$	0.446624
$P(tB\_cont)$	0.795589	$P(tD4)$	0.551712
$P(tE\_B)$	0.040564	$P(tA4)$	0.000145
$P(tL2)$	0.048371	$P(tB4)$	0.001002
$P(tE\_L)$	0.026599	$P(tE\_D4)$	0.000051
$P(tL\_cont)$	0.031718	$P(tML4)$	0.001809
$P(tML)$	0.941683	$P(tD4\_cont)$	0.996992
$P(tA1)$	0.000103	$P(tE\_ML)$	0.003041
$P(tB1)$	0.000708	$P(tML\_cont)$	0.107661
$P(tE\_D1)$	0.000036	$P(tC)$	0.889298
$P(tL1)$	0.000043	$P(tA5)$	0.008104
$P(tML1)$	0.001277	$P(tB5)$	0.055831
$P(tD1\_cont)$	0.000871	$P(tE\_C)$	0.002847
$P(tD2)$	0.996963	$P(tML5)$	0.100778
$P(tA2)$	0.000065	$P(C\_cont)$	0.832441
$P(tB2)$	0.000451		
$P(tE\_D2)$	0.000023		
$P(tML2)$	0.000814		
$P(tD2\_cont)$	0.635502		

Веројатносните дистрибуции на стабилните состојби во соодветниот СТМС, се пресметуваат со примена на infinitesimal generator матрица и системот линеарни равенки, прикажани во поглавје 2.3, чекор 2 и чекор 3, соодветно. Решената infinitesimal generator матрица според студијата на случај за првиот кластер е прикажана во Табела 23.

**Табела 23 Infinitesimal generator матрица за Кластер 1**

<b>Infinitesimal generator matrix</b>									
-0.000139	0.000059	0.000003	0.000004	0.000000	0.000000	0.000073	0.000000	0.000000	0.000000
0.000009	-0.000015	0.000003	0.000004	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000003	-0.000110	0.000107	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000003	0.000000	-0.000885	0.000882	0.000000	0.000000	0.000000	0.000000
0.000009	0.000059	0.000003	0.000000	0.000107	-0.000178	0.000000	0.000000	0.000000	0.000000
0.000009	0.000059	0.000003	0.000004	0.000107	0.000000	-0.083514	0.083333	0.000000	0.000000
0.000009	0.000059	0.000003	0.000000	0.000107	0.000000	0.000000	-0.047797	0.047619	0.000000
0.000009	0.000059	0.000003	0.000000	0.000107	0.000000	0.000000	0.000000	-0.059001	0.058824
0.000009	0.000059	0.000003	0.000000	0.000107	0.000000	0.000000	0.000000	0.000000	-0.000178

Од ова може да се заклучи дека  $\pi_0 = \pi_1 = \pi_3 = \pi_4 = \pi_5 = \pi_6 = \pi_7 = \pi_8 = \pi_9 = 0$  и  $\pi_2 = 1$ , бидејќи состојбата  $M_E$  е апсорбирачка, што значи пресметката на веројатносните дистрибуции на стабилните состојби е безначајна. Сите други мерки од интерес се пресметани и прикажани подолу.

Просечните времиња на престој за допирливите состојби во првиот кластер се прикажани во Табела 24, пресметани според формула ( 2.6 ) од поглавје 2.3.

**Табела 24 Просечно време на престој во состојбите за Кластер 1**

Marking	Time [s]	Time [min]
$M_A$	6796.70	113.28
$M_B$	13451.82	224.20
$M_E$	$\infty$	$\infty$
$M_L$	8820.74	147.01
$M_{ML}$	1008.46	16.81
$M_C$	943.99	15.73
$M_{D1}$	11.96	0.20
$M_{D2}$	7.63	0.13
$M_{D3}$	9.38	0.16
$M_{D4}$	16.95	0.28

Исто така вкупното време на престој во допирливите состојби може да биде пресметано според формула ( 2.8 ), од поглавје 2.3. Ограничувањето на infinitesimal generator матрицата е прикажана во Табела 25.

**Табела 25 Ограничување на Infinitesimal generator матрица за Кластер 1**

Restriction of Infinitesimal generator matrix								
-0.000139	0.000059	0.000004	0.000000	0.000000	0.000073	0.000000	0.000000	0.000000
0.000009	-0.000015	0.000004	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	-0.000110	0.000107	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	-0.000885	0.000882	0.000000	0.000000	0.000000	0.000000
0.000009	0.000059	0.000000	0.000107	-0.000178	0.000000	0.000000	0.000000	0.000000
0.000009	0.000059	0.000004	0.000107	0.000000	-0.083514	0.083333	0.000000	0.000000
0.000009	0.000059	0.000000	0.000107	0.000000	0.000000	-0.047797	0.047619	0.000000
0.000009	0.000059	0.000000	0.000107	0.000000	0.000000	0.000000	-0.059001	0.058824
0.000009	0.000059	0.000000	0.000107	0.000000	0.000000	0.000000	0.000000	-0.000178

Решениот вектор на очекувано време поминато во сите девет преодни допирливи означувања,  $\vec{x}$ , е прикажан во Табела 26.

**Табела 26 Очекувано време поминато во сите преодни допирливи означувања за Кластер 1**

Expected time spent in a transient tangible marking									
$M_A$	$M_B$	$M_L$	$M_{ML}$	$M_C$	$M_{D1}$	$M_{D2}$	$M_{D3}$	$M_{D4}$	
32577.45	338537.98	13496.14	8099.07	40131.37	28.48	49.65	40.07	13241.97	[s]
542.96	5642.30	224.94	134.98	668.86	0.47	0.83	0.67	220.70	[min]

Просечниот број на посети во сите девет преодни состојби,  $n_E$ , е прикажана во Табела 28. Оваа вредност се добива со помош на формула ( 2.10 ) од поглавје 2.3. Ограничувањето на транзициската веројатносна матрица,  $P_E^*$ , е прикажано во Табела 27.

**Табела 27 Ограничување на веројатносната транзициска матрица за Кластер 1**

<b>Restriction of the transition probability matrix</b>								
0.000000	0.426888	0.025955	0.000000	0.000000	0.525392	0.000000	0.000000	0.000000
0.564919	0.000000	0.236638	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.972530	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.996592	0.000000	0.000000	0.000000	0.000000
0.048362	0.333201	0.000000	0.601448	0.000000	0.000000	0.000000	0.000000	0.000000
0.000103	0.000708	0.000043	0.001278	0.000000	0.000000	0.997832	0.000000	0.000000
0.000180	0.001237	0.000000	0.002234	0.000000	0.000000	0.000000	0.996286	0.000000
0.000145	0.001002	0.000000	0.001809	0.000000	0.000000	0.000000	0.000000	0.996992
0.048362	0.333201	0.000000	0.601448	0.000000	0.000000	0.000000	0.000000	0.000000

**Табела 28 Просечен број на посети во сите преодни состојби во Кластер 1**

<b>Total number of visits in a transient tangible marking</b>		
$M_A$	$n_0$	3.501551
$M_B$	$n_1$	3.828545
$M_L$	$n_2$	0.996941
$M_{ML}$	$n_3$	5.182163
$M_C$	$n_4$	5.164502
$M_{D1}$	$n_5$	1.839687
$M_{D2}$	$n_6$	1.835698
$M_{D3}$	$n_7$	1.828880
$M_{D4}$	$n_8$	1.823379

Кумулативното време на престој во сите девет преодни означувања е пресметано според формула ( 2.9 ) од поглавје 2.3, и е прикажано во Табела 29.

**Табела 29 Кумулативното време на престој во сите преодни означувања во Кластер 1**

<b>Cumulative sojourn time in a transient tangible marking</b>								
$M_A$	$M_B$	$M_L$	$M_{ML}$	$M_C$	$M_{D1}$	$M_{D2}$	$M_{D3}$	$M_{D4}$
23798.99	51500.88	8793.76	5226.02	4875.23	22.01	14.00	17.15	30.90
396.65	858.35	146.56	87.10	81.25	0.37	0.23	0.29	0.52

*[s]*

*[min]*

Оттука следува дека, просечното време на траење на единечна корисничка посета во првиот кластер е 94278.95 [s] или 1571.32 [min].

#### 4.2.4.3.2 Резултати за Кластер 2

Пресметаните интензитети на сите временски премини во изградената GSPN мрежа, кои се однесуваат на вториот кластер, се прикажани во Табела 30.

**Табела 30 Интензитети на временските премини пресметани за Кластер 2**

Rate	
$\alpha$	0.000033
$\lambda$	0.008621
$\mu$	0.000003
$\kappa$	0.000003
$\nu$	0.000080
$\theta$	0.001027
$\varepsilon$	0.000208
$\gamma$	0.050000
$\beta$	0.055556

Веројатностите за активирање на временските премини се прикажани во Табела 31:

**Табела 31 Веројатности на активација на премините во Кластер 1**

Probabilities			
$P(tA\_cont)$	0.003774	$P(tD3)$	0.000450
$P(tB)$	0.972033	$P(tA3)$	0.115879
$P(tE\_A)$	0.000340	$P(tB3)$	0.000041
$P(tL)$	0.000377	$P(tE\_D3)$	0.001075
$P(tD1)$	0.023476	$P(tML3)$	0.746777
$P(tA6)$	0.003865	$P(tD3\_cont)$	0.135778
$P(tB\_cont)$	0.995401	$P(tD4)$	0.001777
$P(tE\_B)$	0.000348	$P(tA4)$	0.457618
$P(tL2)$	0.000386	$P(tB4)$	0.000160
$P(tE\_L)$	0.034926	$P(tE\_D4)$	0.004246
$P(tL\_cont)$	0.038736	$P(tML4)$	0.536199
$P(tML)$	0.926338	$P(tD4\_cont)$	0.000450
$P(tA1)$	0.000568	$P(tE\_ML)$	0.002717
$P(tB1)$	0.146241	$P(tML\_cont)$	0.072075
$P(tE\_D1)$	0.000051	$P(tC)$	0.925208
$P(tL1)$	0.000057	$P(tA5)$	0.003428
$P(tML1)$	0.001357	$P(tB5)$	0.882919
$P(tD1\_cont)$	0.003532	$P(tE\_C)$	0.000309
$P(tD2)$	0.848195	$P(tML5)$	0.008192
$P(tA2)$	0.000293	$P(C\_cont)$	0.105153
$P(tB2)$	0.075426		
$P(tE\_D2)$	0.000026		
$P(tML2)$	0.000700		
$P(tD2\_cont)$	0.437473		

Веројатносните дистрибуции на стабилните состојби во соодветниот СТМС, се пресметуваат со примена на infinitesimal generator матрица и системот линеарни равенки, прикажани во поглавје 2.3, чекор 2 и чекор 3, соодветно. Решената infinitesimal generator матрица според студијата на случај за вториот кластер е прикажана во Табела 32.

**Табела 32 Infinitesimal generator матрица за Кластер 2**

Infinitesimal generator matrix									
-0.008835	0.008621	0.000003	0.000003	0.000000	0.000000	0.000208	0.000000	0.000000	0.000000
0.000033	-0.000040	0.000003	0.000003	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000003	-0.000083	0.000080	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000003	0.000000	-0.001030	0.001027	0.000000	0.000000	0.000000	0.000000
0.000033	0.008621	0.000003	0.000000	0.000080	-0.008737	0.000000	0.000000	0.000000	0.000000
0.000033	0.008621	0.000003	0.000003	0.000080	0.000000	-0.058740	0.050000	0.000000	0.000000
0.000033	0.008621	0.000003	0.000000	0.000080	0.000000	0.000000	-0.064293	0.055556	0.000000
0.000033	0.008621	0.000003	0.000000	0.000080	0.000000	0.000000	0.000000	-0.018838	0.010101
0.000033	0.008621	0.000003	0.000000	0.000080	0.000000	0.000000	0.000000	0.000000	-0.008737

И во овој случај може да се заклучи дека  $\pi_0 = \pi_1 = \pi_3 = \pi_4 = \pi_5 = \pi_6 = \pi_7 = \pi_8 = \pi_9 = 0$  и  $\pi_2 = 1$ , бидејќи состојбата  $M_E$  е апсорбирачка, што значи пресметката на веројатносните дистрибуции на стабилните состојби е безначајна. Сите други мерки од интерес се пресметани и прикажани подолу.

Просечните времиња на престој за допирливите состојби во вториот кластер се прикажани во Табела 33, пресметани според формула ( 2.6 ) од поглавје 2.3.

**Табела 33 Просечно време на престој во состојбите за Кластер 2**

Marking	Time [s]	Time [min]
$M_A$	112.76	1.88
$M_B$	115.47	1.92
$M_E$	$\infty$	$\infty$
$M_L$	11582.01	193.03
$M_{ML}$	901.15	15.02
$M_C$	102.42	1.71
$M_{D1}$	16.96	0.28
$M_{D2}$	8.75	0.15
$M_{D3}$	13.44	0.22
$M_{D4}$	53.08	0.88

Исто така вкупното време на престој во допирливите состојби може да биде пресметано според формула ( 2.8 ), од поглавје 2.3. Ограничувањето на infinitesimal generator матрицата е прикажано во Табела 34.



**Табела 34 Ограничување на Infinitesimal generator матрица за Кластер 1**

<b>Restriction of Infinitesimal generator matrix</b>								
-0.008835	0.008621	0.000003	0.000000	0.000000	0.000208	0.000000	0.000000	0.000000
0.000033	-0.000040	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	-0.000083	0.000080	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	-0.001030	0.001027	0.000000	0.000000	0.000000	0.000000
0.000033	0.008621	0.000000	0.000080	-0.008737	0.000000	0.000000	0.000000	0.000000
0.000033	0.008621	0.000003	0.000080	0.000000	-0.058740	0.050000	0.000000	0.000000
0.000033	0.008621	0.000000	0.000080	0.000000	0.000000	-0.064293	0.055556	0.000000
0.000033	0.008621	0.000000	0.000080	0.000000	0.000000	0.000000	-0.018838	0.010101
0.000033	0.008621	0.000000	0.000080	0.000000	0.000000	0.000000	0.000000	-0.008737

Решениот вектор на очекувано време поминато во сите девет преодни допирливи означувања во вториот кластер,  $\vec{x}$ , е прикажан во Табела 35.

**Табела 35 Очекувано време поминато во сите преодни допирливи означувања за Кластер 2**

<b>Expected time spent in a transient tangible marking</b>									
$M_A$	$M_B$	$M_L$	$M_{ML}$	$M_C$	$M_{D1}$	$M_{D2}$	$M_{D3}$	$M_{D4}$	
645.74	142562.32	23.42	3.04	0.36	2.29	1.78	5.24	6.06	[s]
10.76	2376.04	0.39	0.05	0.01	0.04	0.03	0.09	0.10	[min]

Просечниот број на посети во сите девет преодни состојби во рамки на вториот кластер,  $n_E$ , е прикажана во Табела 37. Оваа вредност се добива со помош на формула ( 2.10 ) од поглавје 2.3. Ограничувањето на транзициската веројатносна матрица,  $P_E^*$ , е прикажано во Табела 36.

**Табела 36 Ограничување на транзициската веројатносна матрица за Кластер 2**

<b>Restriction of the transition probability matrix</b>								
0.000000	0.975715	0.000379	0.000000	0.000000	0.023565	0.000000	0.000000	0.000000
0.840319	0.000000	0.083971	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.963667	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.997072	0.000000	0.000000	0.000000	0.000000
0.003831	0.986670	0.000000	0.009154	0.000000	0.000000	0.000000	0.000000	0.000000
0.000570	0.146759	0.000057	0.001362	0.000000	0.000000	0.851201	0.000000	0.000000
0.000521	0.134085	0.000000	0.001244	0.000000	0.000000	0.000000	0.864103	0.000000
0.001777	0.457618	0.000000	0.004246	0.000000	0.000000	0.000000	0.000000	0.536199
0.003831	0.986670	0.000000	0.009154	0.000000	0.000000	0.000000	0.000000	0.000000

**Табела 37 Просечен број на посети во сите преодни состојби во Кластер 2**

Total number of visits in a transient tangible marking		
$M_A$	$n_0$	11.593189
$M_B$	$n_1$	12.600198
$M_L$	$n_2$	1.062461
$M_{ML}$	$n_3$	1.035813
$M_C$	$n_4$	1.032780
$M_{D1}$	$n_5$	0.273193
$M_{D2}$	$n_6$	0.232543
$M_{D3}$	$n_7$	0.200941
$M_{D4}$	$n_8$	0.107744

Кумулативното време на престој во сите девет преодни означувања во вториот кластер е пресметано според формула ( 2.9 ) од поглавје 2.3, и е прикажано во Табела 38.

**Табела 38 Кумулативното време на престој во сите преодни означувања во Кластер 2**

Cumulative sojourn time in a transient tangible marking								
$M_A$	$M_B$	$M_L$	$M_{ML}$	$M_C$	$M_{D1}$	$M_{D2}$	$M_{D3}$	$M_{D4}$
1307.20	1454.90	12305.43	933.43	105.78	4.63	2.03	2.70	5.72 [s]
21.79	24.25	205.09	15.56	1.76	0.08	0.03	0.05	0.10 [min]

Оттука следува дека, просечното време на траење на единечна корисничка посета во вториот кластер е 16121.82 [s] или 268.70 [min].

#### 4.2.5 Завршен преглед на поглавјето

Во ова поглавје предложен е GSPN пристап за моделирање и опишување на корисничкото однесување при процесите на пребарување во контекст на реформулирање на прашањето [50] и навигација низ информациска мрежа, илустрирано низ примерот на електронска книжарница (продавница) [3]. Во оваа класа на Петриеви мрежи, моменталните премини имаат тежини на активација, временските транзиции имаат експоненцијално дистрибуирани времиња на активација, во чија основа се наоѓа стохастичкиот процес на Марков Ланец во Континуирано Време. За реализација на истражувањето во оваа насока, во делот наменет на процесот на пребарување, се креира и процесира лог датотека на клиентска страна, така што сите идентификувани акции на реформулирање на прашање во рамки на процесот на пребарување се претставени како премини. Исто така, во делот наменет на процесот на навигација целокупните анализи беа извршени врз навигациските податоци собирани на серверска страна во одредена временска рамка и се однесуваат на историски податоци за страни до кои е пристапено во рамки на електронската книжарница. И во овој случај, сите идентификувани навигациски акции, се претставени како премини. Со цел евалуација на мерките на перформанси, беше применет ефикасен алгоритам за пресметка на стабилните состојби,

предложен од [17]. Добиените резултати се однесуваат на: просечно време на престој за секоја од преодните состојби, вкупно време поминато во овие преодни состојби, просечен број на посети, како и кумулативно време на престој [19].

#### 4.2.6 Идни правци

Предложениот бихејвиористички-базиран модел на процесот на пребарување, освен што дава визуелна претстава за шемите на однесување на корисниците при процесот на пребарување во контекст на реформулирање на прашањето, во иднина може да се искористи и прилагоди за развој на систем кој ќе поддржува препораки на прашања или персонализирано пребарување, од различни аспекти, како на пример: историја на претходно поставени прашања, откриени интереси, намери и цели [6], [7], [26], [34], [62], [64], и сл.

Исто така, предложениот бихејвиористички-базиран модел на процесот на навигација, може да се насочи кон: (1) од една страна развој на систем за препораки, пред сè во сферата на електронската трговија, со цел: персонализација на процесот на купување, зголемување на продажбата и претворање на што поголем дел пребарувачи (browsers) во купувачи (buyers), забрзување и зголемување на можноста на избор од што поголемо множество продукти задоволувајќи ги желбите на купувачите, задржување на вниманието на купувачите, нивната лојалност и сл. [33], [36], [39], [65], [66], [67], [74], [83]; (2) а од друга страна, прилагодување и подобрување на веб дизајнот, односно интерфејсот, според интересите и однесувањето на корисниците при процесот на навигација со цел подобрување на нивното искуство, како и подобрување на интеракцијата со системот, повторно со посебен акцент на домаинот на електронската трговија [38], [82]. Двата модели може да се соединат со цел да се добие една детална и поопширна слика за шемите на однесување на корисниците и нивните интереси во една ваква околина.

На крај, за мерење на ефективноста на електронската продавница во остварувањето на своите бизнис цели, интересно би било да се придонесе во пресметката на бизнис-ориентираните мерки на перформанси, како: Buy-to-visit ratio, Revenue throughput (\$/sec), Potential loss throughput or Lost revenue throughput (\$/sec), за што, повеќе детали можат да се најдат во [42].

Следното поглавје е наменето да се види како развиените модели од ваков тип може да се искористат за мерење ангажираност на корисникот. Пред да се премине на тоа, ќе биде даден детален опис, дискусија и пресметка на некои компоненти на ангажираност на корисникот преку искористување на квантитативните податоци и извлечените статистички заклучоци.

## ***Глава 5***

## 5 Прилог кон пресметка на компонентите на ангажираност на корисникот

---

Освен направената евалуацијата на мерките за мерење на перформанси за двата предложени бихејвиористички-базирани модели, повторно во насока на корисничкото искуство, интересно би било за прв пат да се види како модели од ваков тип може да се искористат за мерење ангажираност на корисникот. Во таа насока, ова поглавје претставува прилог кон секција 3.4. Пред да се премини на тој аспект, преку искористување на квантитативните податоци, односно лог податоците, и извлечените статистички заклучоци, во продолжение ќе биде даден детален опис, дискусија и пресметка на некои компоненти на ангажираност на корисникот.

Во моделот на Петерсон [51], ангажираноста на корисникот е дефинирана како мерка која се пресметува од 7 различни мерки (индекси), преку следната равенка:

$$UE = \sum (C_i + D_i + R_i + L_i + B_i + F_i + I_i) \quad (5.1)$$

Првата компонента претставува ClickDepth индекс ( $C_i$ ) и се однесува на процентот на корисници кои имаат број на прегледи на страни по сесија поголем или еднаков на некоја утврдена гранична вредност. На ниво на корисник, овој индекс се пресметува според следната формула:

$$C_i = \text{Sessions having at least "threshold" page views} / \text{All sessions} \quad (5.2)$$

Втората компонента претставува Duration индекс ( $D_i$ ) и се однесува на процентот на корисници чие време на престој по сесија е поголемо или еднакво на некоја утврдена гранична вредност. На ниво на корисник, оваа мерка се пресметува според следната формула:

$$D_i = \text{Sessions having duration over "threshold" time} / \text{All sessions} \quad (5.3)$$

Третата компонента или Resency индекс ( $R_i$ ) го претставува процентот на корисници кои доаѓаат и заминуваат од веб сајтот во границите на некоја предодредена временска рамка, најчесто изразена преку денови. Со други зборови, овој индекс се однесува на честотата на прегледување на веб сајтот. Колку посетите (сесиите) се почести, толку вредноста на индексот е поголема. За секој корисник поединечно, индексот се пресметува според следната формула:

$$R_i = 1 / \text{Number of days elapsed since the most recent session} \quad (5.4)$$

Четвртата компонента претставува Loyalty индекс ( $L_i$ ) кој се однесува на процентот на корисници чија фреквенција на повторувачки посети (сесии) е еднаква или поголема од

некоја претходно утврдена гранична вредност. Основната формула на ниво на корисник е следната:

$$L_i = 1 - (1/\text{Number of visitor sessions during the time frame}) \quad (5.5)$$

Петтата компонента или Brand индекс ( $B_i$ ) го претставува процентот на корисници кои доаѓаат на веб сајтот директно или преку брендиран пребарувачки термини.

Шестата компонента или Feedback индекс ( $F_i$ ) го претставува процентот на корисници кои оставаат повратна информација на веб сајтот, некаков рејтинг или коментари. На ниво на корисник се пресметува според следната формула:

$$F_i = \text{Sessions where visitor submits feedback} / \text{All sessions} \quad (5.6)$$

Последната седма компонента или Interaction индекс ( $I_i$ ) се однесува на процентот на корисници кои посетуваат, односно имаат интеракција со одредена содржина на веб сајтот која е од големо значење, на пример за електронска продавница, ставање во кошничка или купување на производ. За секој корисник поединечно се пресметува според следната формула:

$$I_i = \text{Sessions where visitor completes an action} / \text{All sessions} \quad (5.7)$$

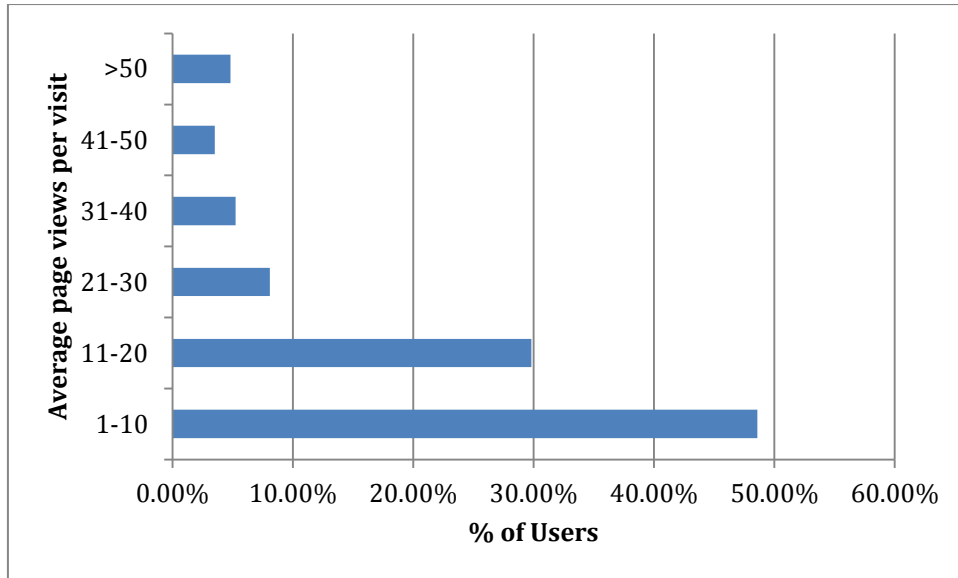
Доколку не постои можност да се пресмета некоја од компонентите на моделот на Петерсон, таа едноставно се заменува со вредност нула, што не влијае на крајниот резултат.

## 5.1 Користење на лог податоците за пресметка на компонентите на ангажираност на корисникот

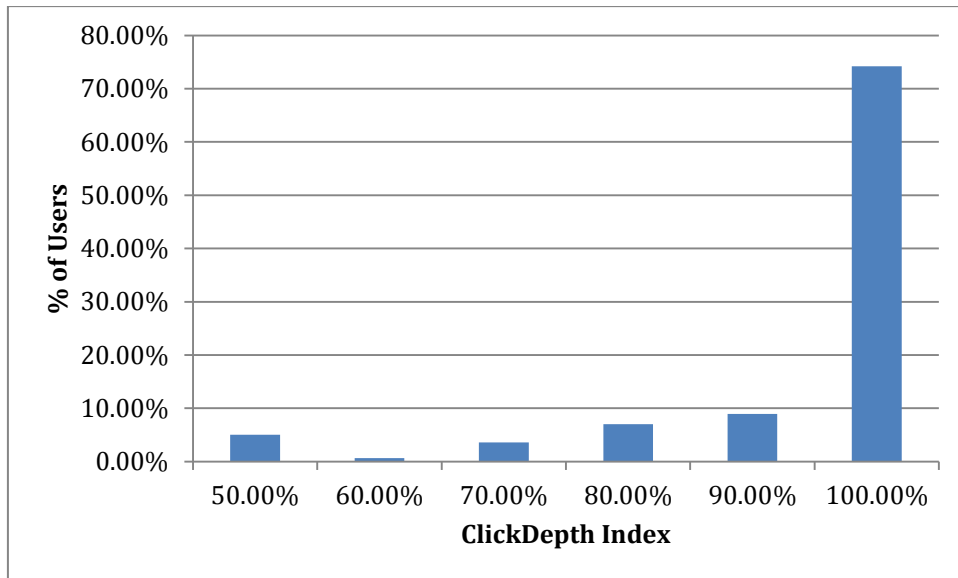
Во оваа докторска дисертација за пресметка на индексите илустрирана е примена на времето помеѓу две последователни посети, наместо оригиналниот пристап каде се применуваат сесии. Од направената статистичка анализа на лог податоците, се утврдија потребните гранични вредности и параметри кои учествуваат во пресметката на индексите.

За првиот кластер:

- За  $C_i$ , беше утврдено дека најголем процент од корисниците (околу 48%) имаат посети каде просечниот број на прегледи на страни по посета е помеѓу 1 и 10. Графички тоа е прикажано на Слика 22. Оваа гранична вредност беше искористена за пресметка на  $C_i$ , така што распределбата на  $C_i$  вредностите за сите 916 корисници и нивните 7570 посети е прикажана на Слика 23.

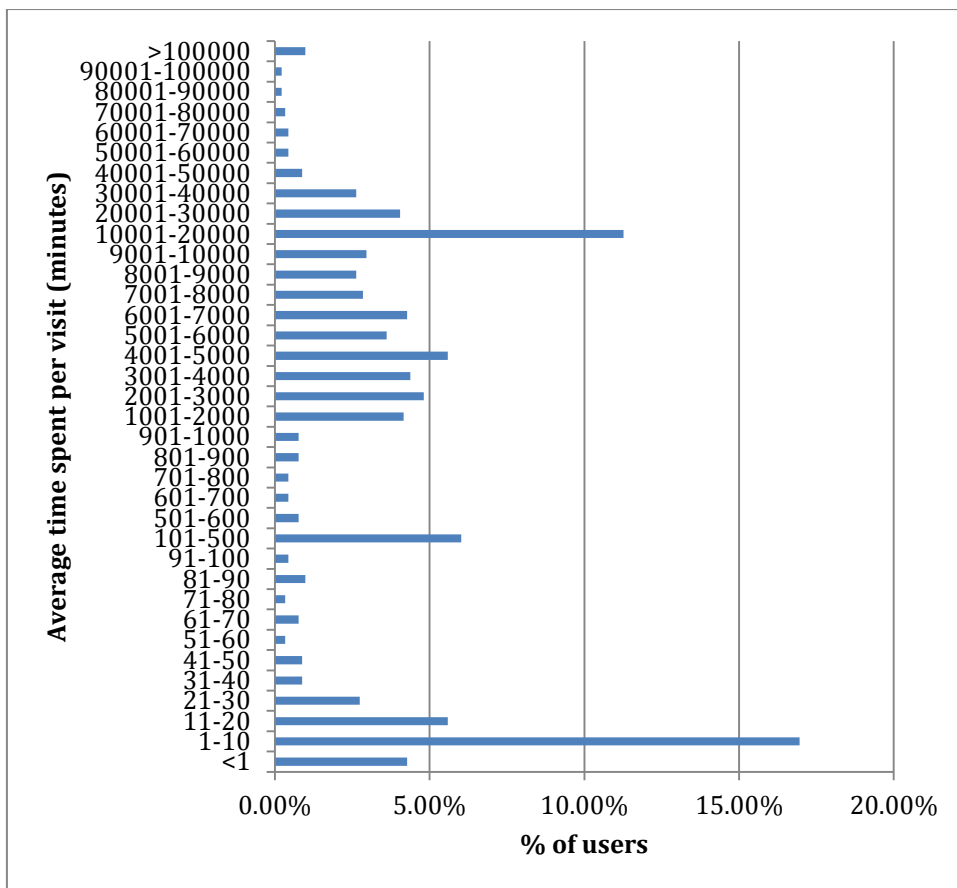


Слика 22 Просечен број на прегледи на страни по посета во Кластер 1

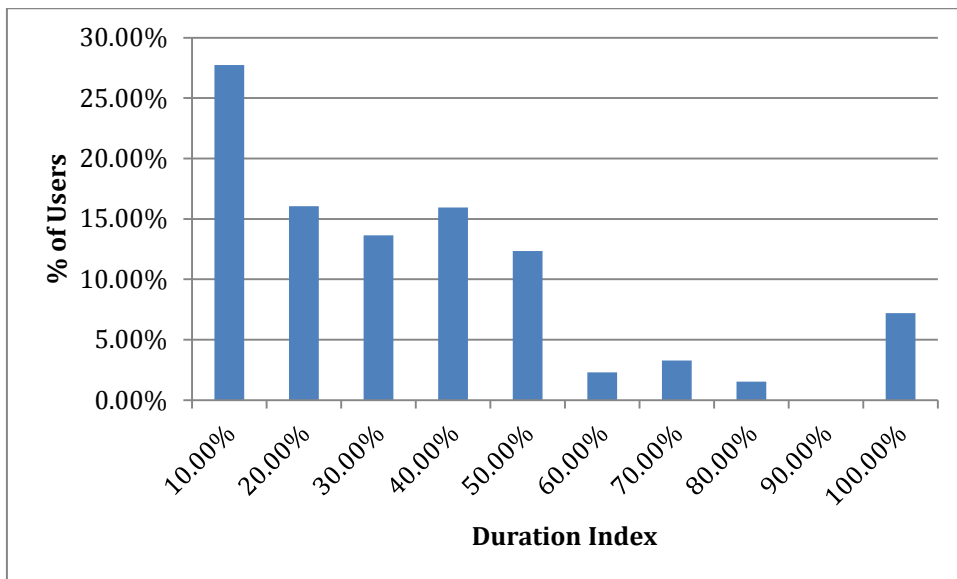


Слика 23 Распределба на  $C_i$  вредностите за 916 корисници и 7570 посети од Кластер 1

- За  $D_i$ , беше утврдено дека најголем процент од корисниците (околу 17%) имаат посети со просечна должина помеѓу 1 и 10 минути. Графички тоа е прикажано на Слика 24 . Оваа гранична вредност беше искористена за пресметка на  $D_i$ , така што распределбата на  $D_i$  вредностите за сите 916 корисници и нивните 7570 посети е прикажана на Слика 25.



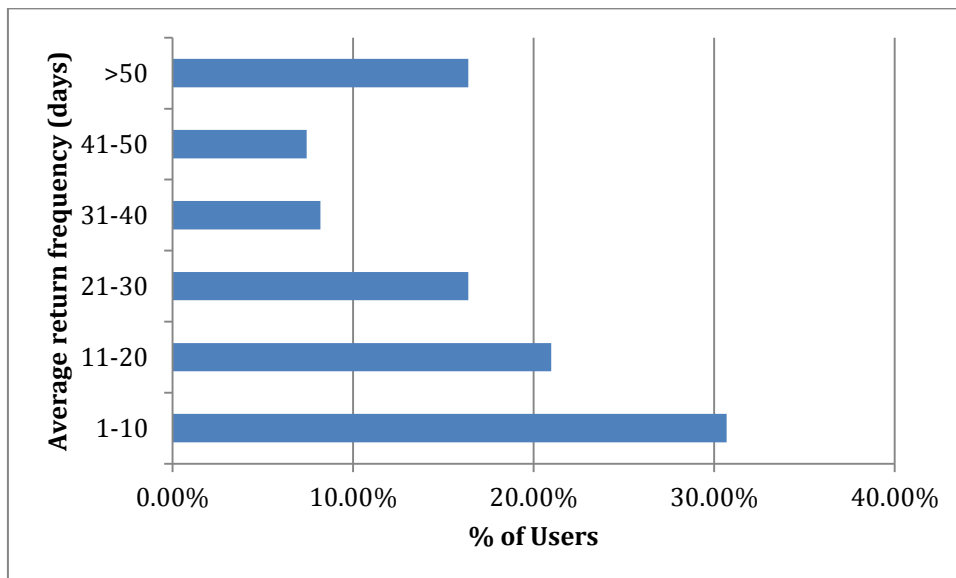
Слика 24 Просечна должина на посетите во Кластер 1



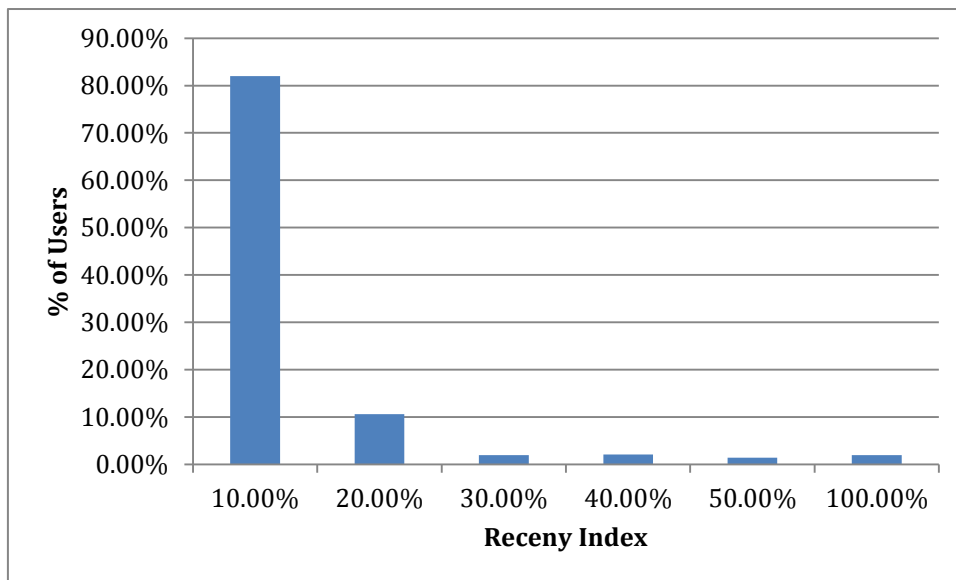
Слика 25 Распределба на  $D_i$  вредностите за 916 корисници и 7570 посети од Кластер 1



- За  $R_i$ , беше утврдено дека најголем процент од корисниците (околу 31%) се враќаат повторно на веб сајтот помеѓу 1 и 10 дена. Графички тоа е прикажано на Слика 26. Оваа вредност беше искористена за пресметка на  $R_i$ , така што распределбата на  $R_i$  вредностите за сите 916 корисници и нивните 7570 посети е прикажана на Слика 27.



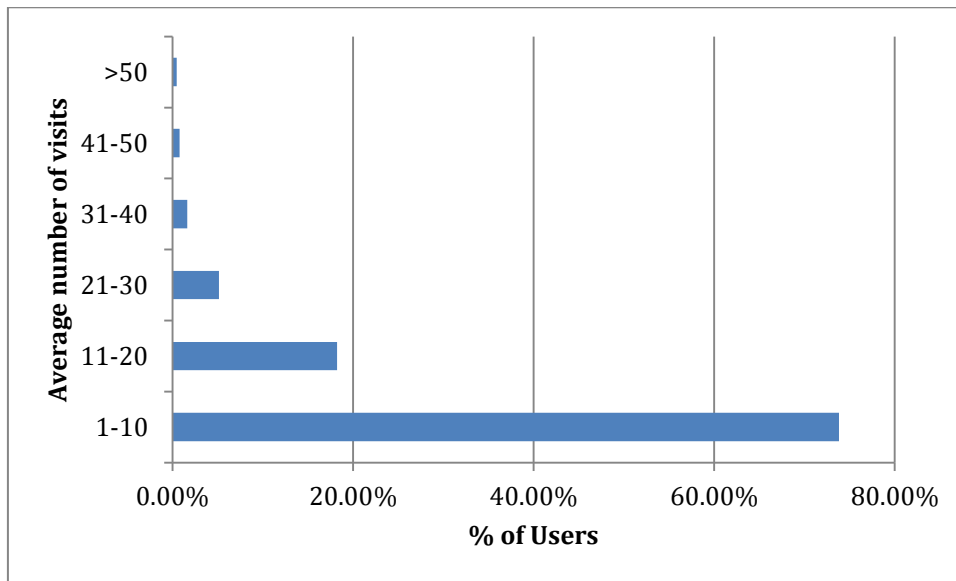
Слика 26 Просечена фреквенција на повторна посета во Кластер 1



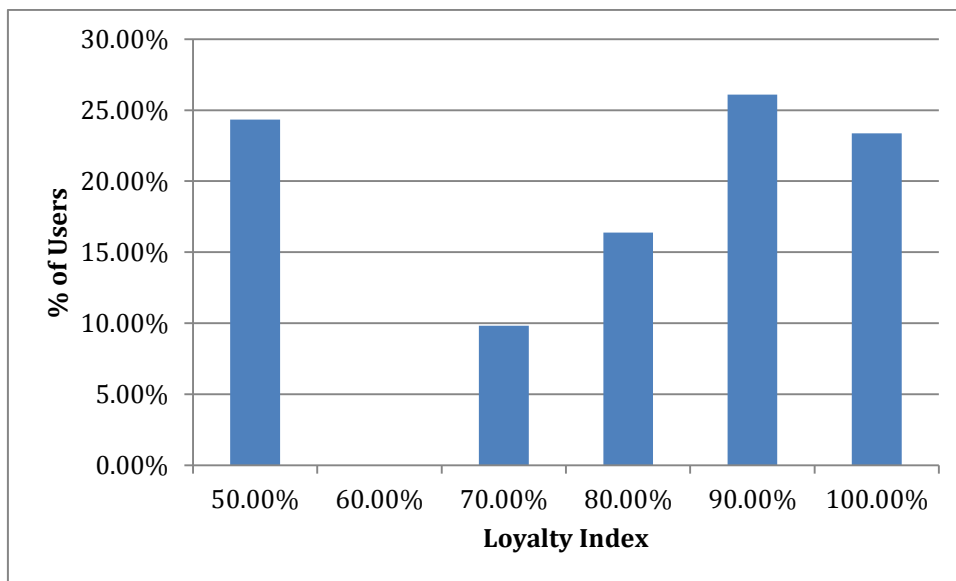
Слика 27 Распределба на  $R_i$  вредностите за 916 корисници и 7570 посети од Кластер 1

- За  $L_i$ , беше утврдено дека најголем процент од корисниците (околу 74%) имаат од 1 до 10 посети на веб сајтот. Графички тоа е прикажано на Слика 28. Оваа вредност беше искористена за пресметка на  $L_i$ , така што распределбата на  $L_i$

вредностите за сите 916 корисници и нивните 7570 посети е прикажана на Слика 29.

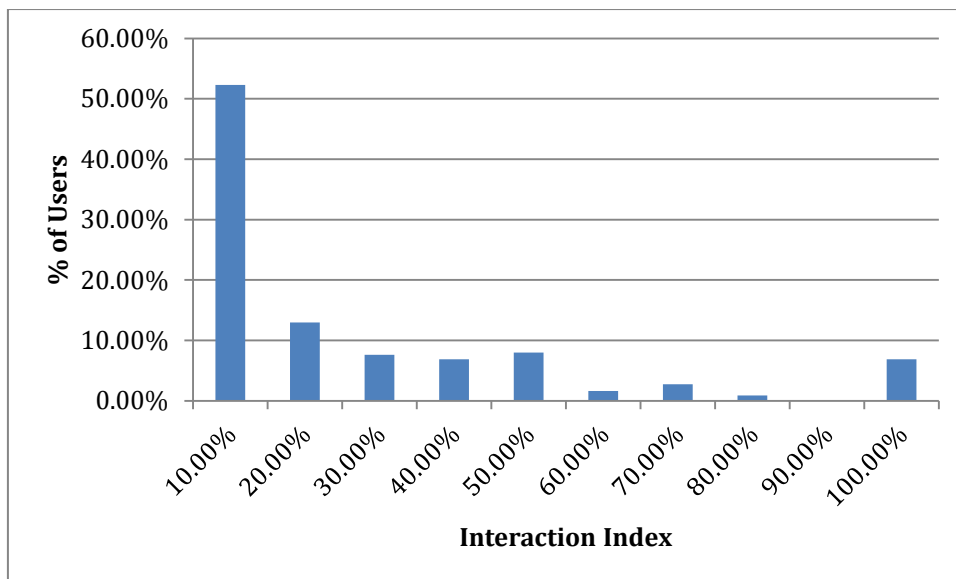


Слика 28 Просечен број посети во Кластер 1



Слика 29 Распределба на  $L_i$  вредностите за 916 корисници и 7570 посети од Кластер 1

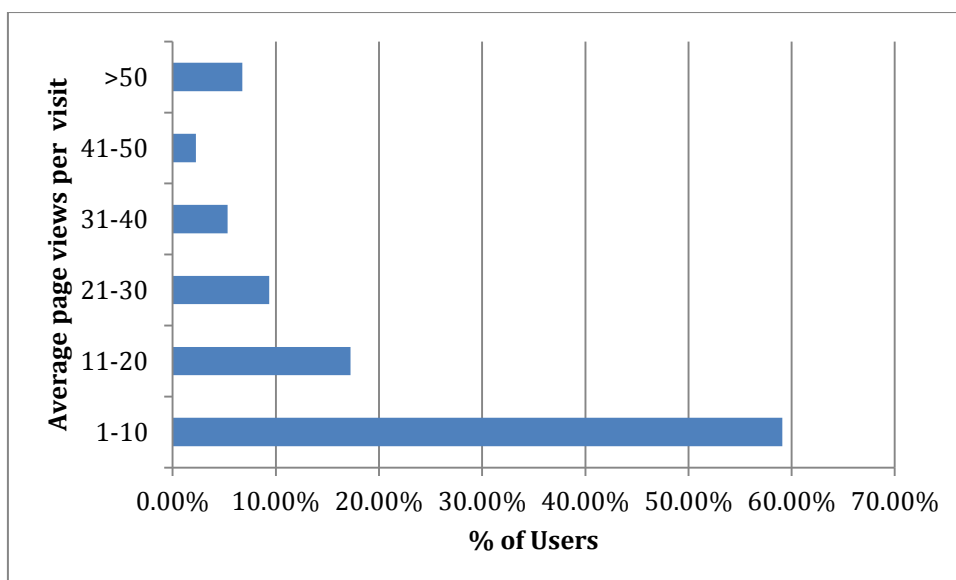
- За пресметка на  $I_i$ , утврдените содржини од интерес за електронската книжарница се ставање на книга во кошничка и купување на книга. Распределбата на  $I_i$  вредностите за сите 916 корисници и нивните 7570 посети е прикажана на Слика 30.



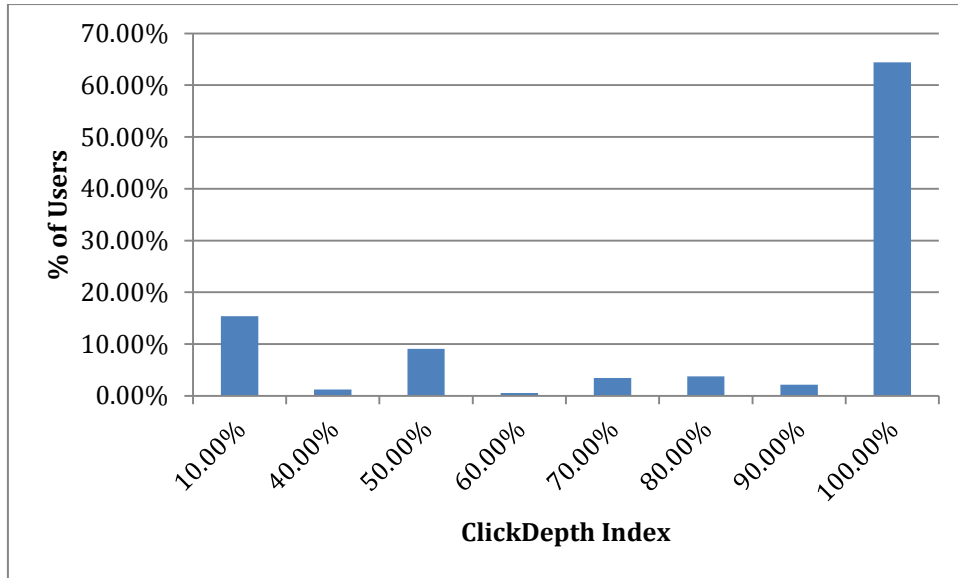
Слика 30 Распределба на  $I_i$  вредностите за 916 корисници и 7570 посети од Кластер 1

За вториот кластер:

- За  $C_i$ , беше утврдено дека најголем процент од корисниците (околу 59%) имаат посети каде просечниот број на прегледи на страни по посета е помеѓу 1 и 10. Графички тоа е прикажано на Слика 31. Оваа гранична вредност беше искористена за пресметка на  $C_i$ , така што распределбата на  $C_i$  вредностите за сите 1068 корисници и нивните 7863 посети е прикажана на Слика 32.

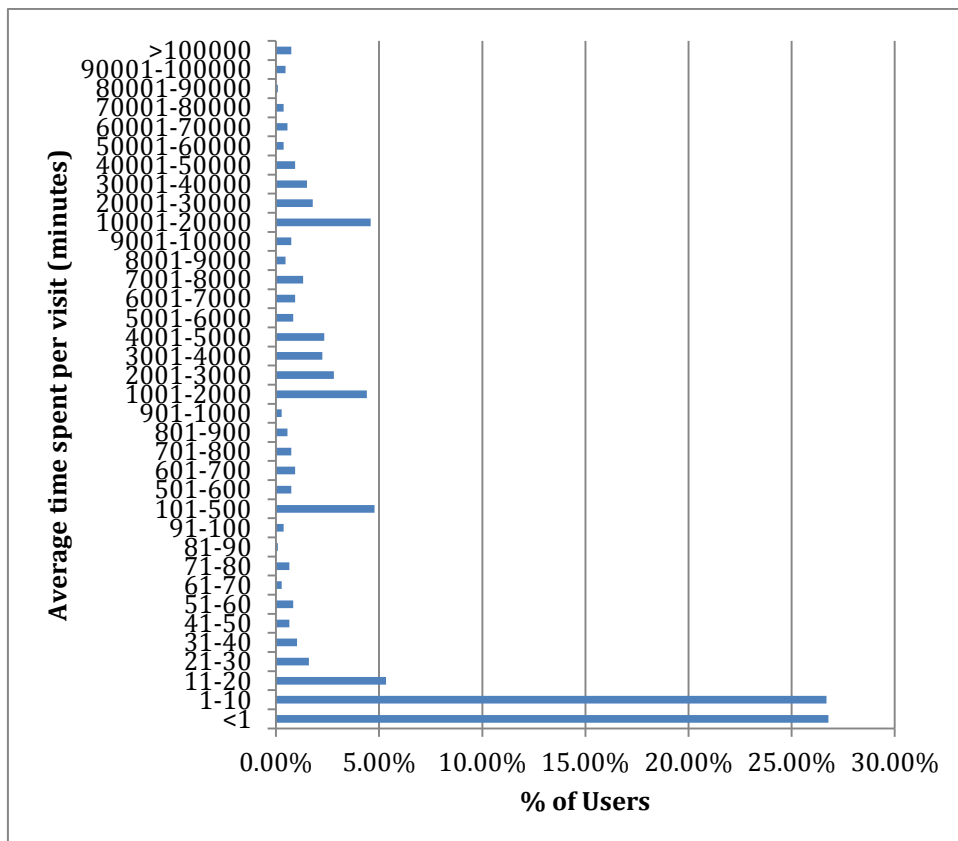


Слика 31 Просечен број на прегледи на страни по посета во Кластер 2

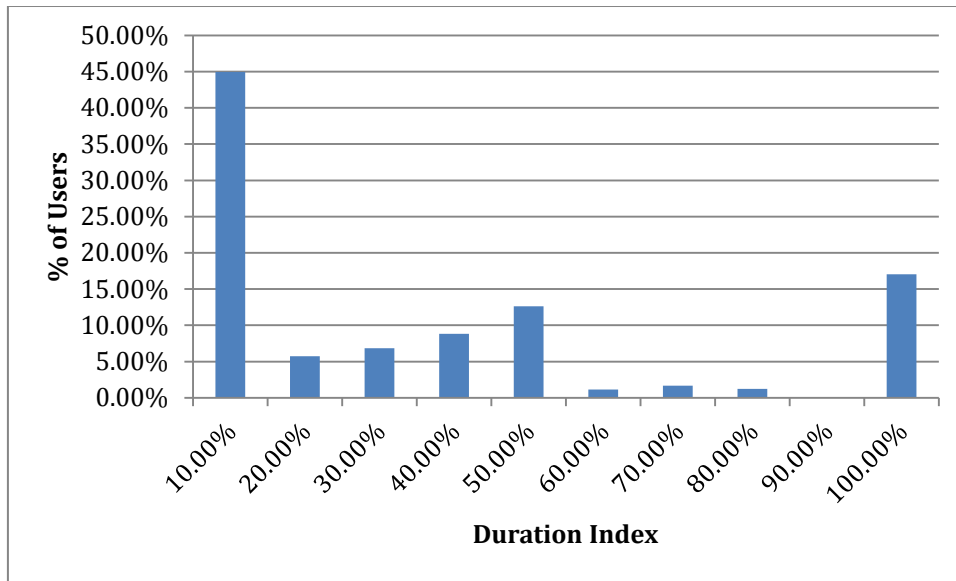


Слика 32 Распределба на  $C_i$  вредностите за 1068 корисници и 7863 посети од Кластер 2

- За  $D_i$ , беше утврдено дека најголем процент од корисниците (околу 27%) имаат посети со просечна должина помала од 1 минута и помеѓу 1 и 10 минути. Графички тоа е прикажано на Слика 33. Која било гранични вредности од овие две да се искористи за пресметка на  $D_i$ , би се добиле скоро исти резултати. Така, распределбата на  $D_i$  вредностите за сите 1068 корисници и нивните 7863 посети е прикажана на Слика 34.

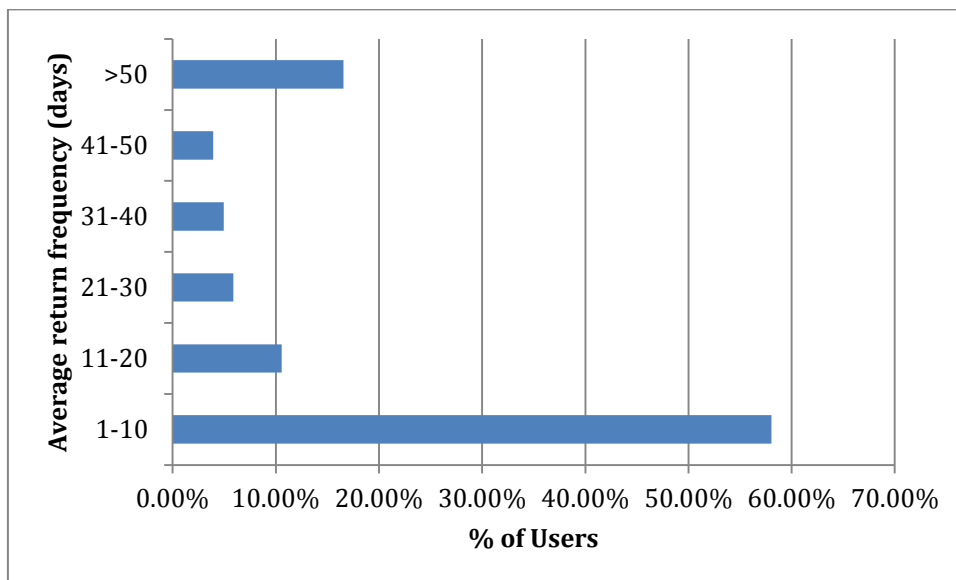


Слика 33 Просечна должина на посетите во Кластер 2

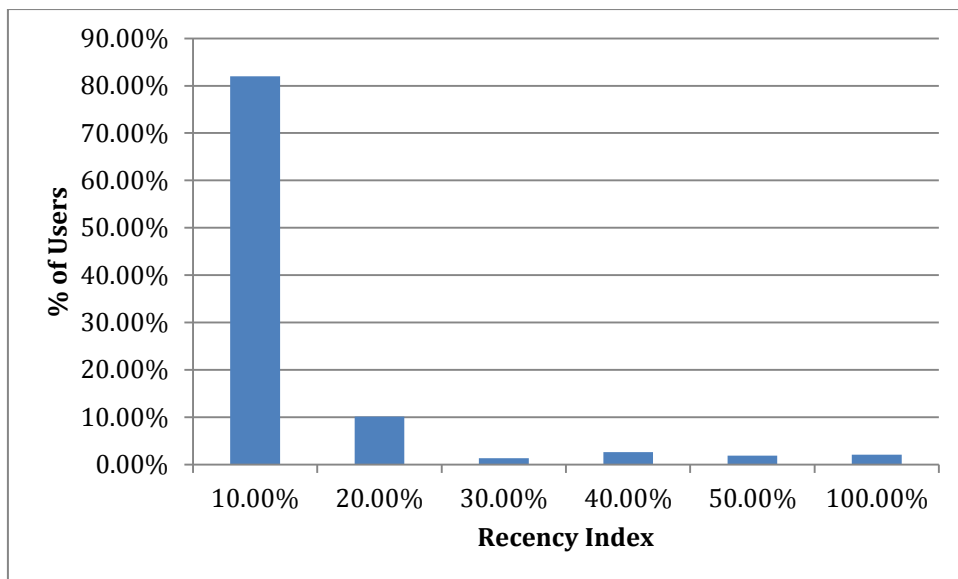


Слика 34 Распределба на  $D_i$  вредностите за 1068 корисници и 7863 посети од Кластер 2

- За  $R_i$ , беше утврдено дека најголем процент од корисниците (околу 58%) се враќаат повторно на веб сајтот помеѓу 1 и 10 дена. Графички тоа е прикажано на Слика 35. Оваа вредност беше искористена за пресметка на  $R_i$ , така што распределбата на  $R_i$  вредностите за сите 1068 корисници и нивните 7863 посети е прикажана на Слика 36.

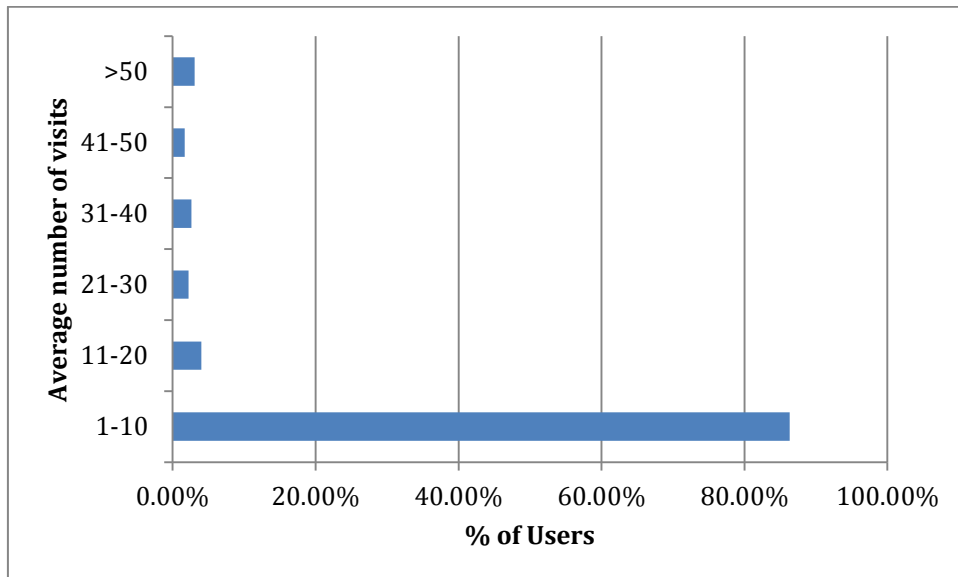


Слика 35 Просечена фреквенција на повторна посета во Кластер 2

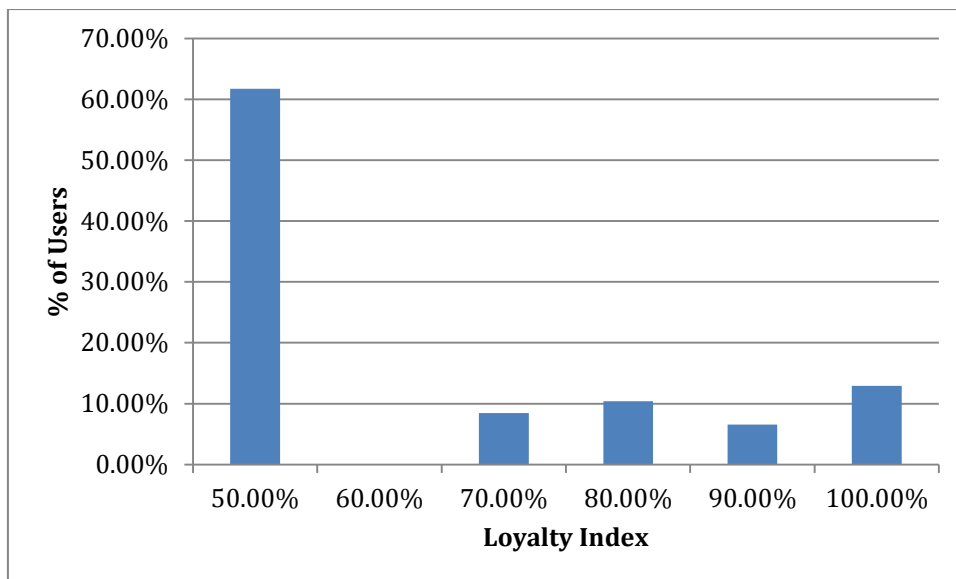


Слика 36 Распределба на  $R_i$  вредностите за 1068 корисници и 7863 посети од Кластер 2

- За  $L_i$ , беше утврдено дека најголем процент од корисниците (околу 86%) имаат од 1 до 10 посети на веб сајтот. Графички тоа е прикажано на Слика 37. Оваа вредност беше искористена за пресметка на  $L_i$ , така што распределбата на  $L_i$  вредностите за сите 1068 корисници и нивните 7863 посети е прикажана на Слика 38.

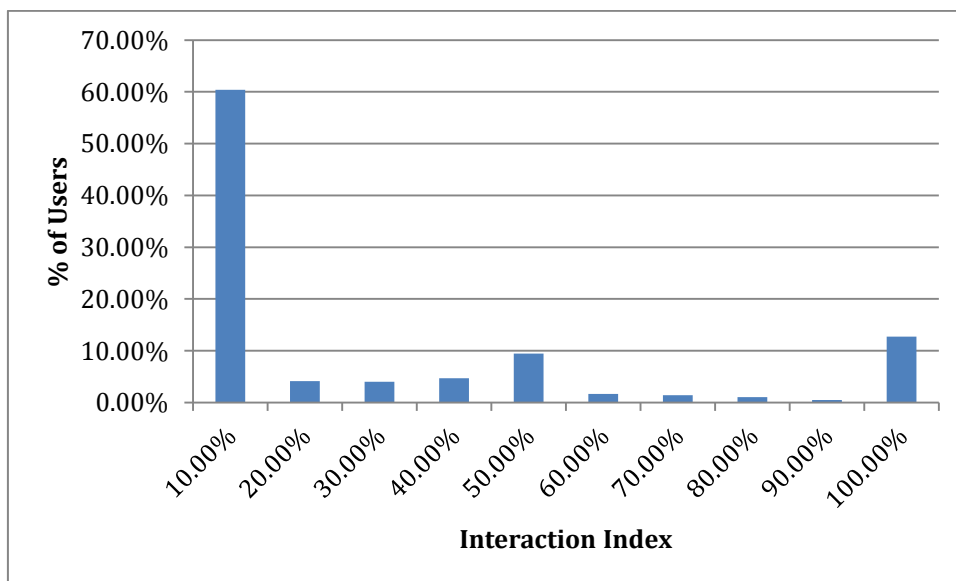


Слика 37 Просечен број посети во Кластер 2



Слика 38 Распределба на  $L_i$  вредностите за 1068 корисници и 7863 посети од Кластер 2

- И овој пат, за пресметка на  $I_i$ , утврдените содржини од интерес за електронската книжарница се ставање на книга во кошничка и купување на книга. Распределбата на  $I_i$  вредностите за сите 1068 корисници и нивните 7863 посети е прикажана на Слика 39.



Слика 39 Распределба на  $I_i$  вредностите за 1068 корисници и 7863 посети од Кластер 2

- Исто така, лог датотеката не содржи информации кои можат да се искористат за пресметка на  $B_i$  и  $F_i$  компонентите. Затоа, истите се заменуваат со вредност 0 во формулата за пресметка на ангажираноста на корисникот.

Следува дека просечната вредност на ангажиранос на корисникот во првиот и вториот кластер е:

$$\begin{aligned} UE_{cluster1} &= (89.80 + 69.24 + 27.93 + 62.62 + 0 + 0 + 38.88)/7 = 41.21\% \\ UE_{cluster2} &= (86.05 + 81.00 + 30.44 + 34.78 + 0 + 0 + 57.24)/7 = 41.36\% \end{aligned} \quad (5.8)$$

Според избраното сценарио и временската рамка во која се собрани лог податоците, можеме да се заклучи дека корисниците од вториот кластер се за нијанса поангажирани во однос на корисниците од првиот кластер.

Исто така, како прилог на овие компоненти, може да се измерат и conversion rate и abandonment rate мерките, кои заедно го отсликуваат корисничкото однесување и укажуваат на корисничкото искуство и ангажиранос на корисникот, така што се добива една подетална слика на целиот овој процес. И двете мерки, може да се пресметаат во однос на целото множество кориснички акции или во однос на различните (уникатните) корисници. Во овој случај, беше избран вториот пристап, така што conversion rate за купување на книга изнесува 18.25%, имајќи во предвид дека во избраната временска рамка електронската книжарница била посетена од 1984 уникатни корисници, од кои само 362 успешно ги завршиле своите трансакции со купување. Тоа значи дека од 961 корисници кои додале книга во својата кошничка, дури една третина станале купувачи, а две третини ја напуштиле својата кошничка без купување, водејќи кон abandonment rate од 62.33%. Исто така, може да се погледне и деталната слика по кластери. За првиот кластер conversion rate изнесува околу 20.21%, што значи дека само мал дел од 916 уникатни корисници успешно ги завршиле своите трансакции со купување книга, што е за нијанса подобро од случајот во вториот кластер каде од 1068 корисници само 16.57% имаат трансакции со купување. Тоа значи дека, слично на глобалната слика, од 514 корисници во првиот кластер, односно 449 корисници од вториот кластер, кои додале книга во својата кошничка, околу една третина станале купувачи, а две третини ја напуштиле својата кошничка без купување, водејќи кон abandonment rate од 64.01% во првиот, односно 60.58% во вториот кластер. Во сите три случаи, abandonment rate е во рамки на усвоените граници, а блиску до моменталниот установен просек<sup>10</sup>.

---

<sup>10</sup> <http://baymard.com/lists/cart-abandonment-rate>



## 5.2 Користење на бихејвиористички-базираните модели за пресметка на компонентите на ангажираност на корисникот

Под претпоставка дека во одредена временска рамка, започнати се повеќе кориснички сесии, односно се направени повеќе кориснички посети, така што следејќи ја “играта” со жетони прикажана на Слика 8, може да се искористи развиениот бихејвиористички-базиран модел на процесот на навигација за пресметка на мерки од ваков тип.

Во оваа секција се предлага начин на пресметка на дел од компонентите на моделот на Петерсон, како и пресметка на conversion rate и abandonment rate.

$C_i$  компонентата може да се пресмета како просечен број на активирања на даден премин во единица време, што уште претставува фреквенција на активирање на преминот [17], [19]. Ако фреквенцијата на преминот е прикажана во формула ( 5.9 ),  $C_i$  може да се изрази преку формула ( 5.10 ). Во суштина, на овој начин се пресметуваат бројот на кликови во единица време една посета, така што секоја посета има различен број на кликови.

$$f_j = \sum_{M_i \in A_j} \lambda_j * \pi_i \quad (5.9)$$

$$C_i = \sum f_j * session\ duration \quad (5.10)$$

Тука,  $M_i$  претставува означување, кое припаѓа на множеството означувања  $A_j$  во кои преминот е овозможен.

Преку развиениот модел може да се пресмета просечното време на престој во преодно допирливо означување, но бидејќи моделот не располага со информации за распределбата на времињата помеѓу две последователни посети, следува дека не може да се пресмета  $D_i$  компонентата. Исто така, не водејќи евиденција во моделот за секој корисник поединечно и неговите посети, не постои можност да се пресметаат компонентите  $R_i$  и  $L_i$ . Исто како во случајот со лог податоците, моделот не е развиен во насока да може да се пресметаат  $B_i$  и  $F_i$  компонентите.

Од друга страна,  $I_i$  компонентата може да се пресмета како тотална веројатност корисникот да извршил одредена акција [17], [19], на пример, во нашиот случај купување книга може да биде акција од интерес. Според моделот, ако состојбата  $D_4$  означува купена книга, до која се стигнува преку секвенцијално активирање на премините  $tD_1 \rightarrow tD_2 \rightarrow tD_3 \rightarrow tD_4$ ,  $I_i$  може да се пресмета како во формула ( 5.11 ):

$$I_i = P(t_{D1}) * P(t_{D2}) * P(t_{D3}) * P(t_{D4}) \quad (5.11)$$

Уште еднаш, за истата посакувана акција, немајќи евиденција за бројот на корисници во моделот, пресметката на conversion rate е изедначена со пресметката на  $I_i$  компонентата, односно терминот корисници е изедначен со терминот време помеѓу две последователни посети, и се пресметува како што е прикажано во формулата ( 5.11 ).

Abandonment rate може да се пресмета според формулата ( 5.12 ), односно колкава е веројатноста корисникот да не купил книга, но да ја ставил во својата кошничка, така што преминот  $tD_1$  означува процес на ставање на книга во кошничка, а преминот  $tD_4$  означува купување на таа книга:

$$abandonment\ rate = (1 - P(tD_4)) * P(tD_1) \quad (5.12)$$

### 5.3 Завршен преглед на поглавјето

Во ова поглавје даден е детален опис на некои компоненти на ангажираност на корисникот, како доста акуелна и атрактивна област на истражување, односно обидување да се изнајдат различни начини и пристапи за зголемување на бројот на ангажирани и задоволни корисници.

Преку набљудување на корисничкото однесување при процесот на навигација, користејќи ги лог податоците, илустрирана е директна пресметка на компонентите на ангажираност на корисникот кои го сочинуваат моделот на Петерсон [51] за двата идентификувани кластери на корисници поединечно. Исто така, во прилог на тоа, пресметани се две дополнителни мерки conversion rate и abandonment rate, со чија помош се доби подетална слика за ангажираноста на корисникот и целокупното корисничко искуство. Покрај употребата на квантитативните податоци, предложена е употреба на развиениот бихејвиористички-базиран модел на навигација за мерење на ангажираноста на корисникот. Сепак, вториот аспект, бара промена и усовршување на моделот, во насока на овозможување пресметка на сите компоненти на ангажираност на корисникот. Ова, може да претставува дополнување на насоките за истражување во иднина, споменати во секција 4.2.6.

Следното поглавје е заклучно поглавје, каде ќе бидат изведени генерални заклучоци за целокупната истражувачка работа.

## ***Глава 6***

## 6 Заклучок

---

Со оваа докторска дисертација направен е обид за продлабочено разбирање на однесувањето на корисниците, генерално при процесот на пребарување и специфично при процесот на навигација, во рамки на информациска мрежа од сферата на електронската трговија, илустрирано низ примерот на првата македонска електронска книжарница – [www.kupikniga.mk](http://www.kupikniga.mk).

Бихејвиористички-базираното моделирање е реализирано со помош на Петриевите мрежи со стохастичко поведение, така што во таа насока, најпрво даден е детален териски приказ на овој концепт, неговите елементи, карактеристики и илустрирано од каде потекнува неговата динамика. Следно, проучено е однесувањето на корисниците при пребарување и навигација, така што предложени се два нови модели кои ги отсликуваат генералните шеми на однесување на корисниците во насока на реформулирање на прашање, од една страна, и однесување на корисниците при навигација, од друга страна. Откако развиените модели беа конструирани и валидирани, пристапено е кон нивно решавање според студија на случај и евалуирање со примена на ефикасен алгоритам во време и простор за пресметување на стабилните состојби, кој е заснован на Марков Ланец во Континуирано Време. Интересен аспект е што моделите за навигација, се решени во насока на два специфични типа на корисници, чии различни шеми на однесување беа автоматски утврдени со помош на алгоритам кој преку комбинација од стандардни методи за кластерирање и техники за анализа на секвенци, засновани на Марков Ланец, групира корисничко однесување според сличноста на секвенцијалниот редослед на кориснички акции. Процесот на евалуирање подразбира, пресметување на одредени мерки на перформанси, какви што се: очекувано време поминато во преодно допирливо означување, кумулативно време на престој во преодно допирливо означување, вкупен број на посети во преодно допирливо означување и сл.

Оваа истражување има директен придонес во анализата на сложениот профил на корисникот и разбирањето на неговото однесување, односно интеракција со системот, со цел да се помогне во подобро прилагодување на сè што претставува хаос, ставајќи го на располагање спектарот на веќе познати шеми на однесување. Сепак, на повидок има бројни отворени предизвици, чиј карактер се менува со исклучиво брзото темпо на менување на трендовите во светот на Интернетот, а пред се во сферата на електронската трговија, што подразбира менување и подобрување на сè што досега е развиено во рамки на оваа докторска дисертација.

Развиените модели може меѓусебно да се комбинираат со цел да се добие една поопширна слика за однесувањето на корисниците и може да се насочи кон, од една страна, персонализирана електронска трговија, а од друга страна, прилагодување и подобрување на веб дизајнот, односно интерфејсот, според интересите и однесувањето на корисниците, а сето тоа со цел подобрување на нивното искуство, зголемување на продажбата, лојалноста и секако задржување на нивното внимание.

На крај, за прв пат, предложена е примена на развиениот GSPN модел за мерење на ангажираност на корисникот, но детална обработка е оставена за во иднина.

## Библиографија

---

- [1] Ajmone Marsan, M., Conte, G. and Balbo, G. (1984) 'A Class of Generalized Stochastic Petri Nets for the Performance Evaluation of Multiprocessor Systems', *ACM Transactions on Computer Systems*, Vol. 2, No. 2, pp. 93-122.
- [2] Al-Maskari, A. and Sanderson, M. (2006) 'The Effects of Topic Familiarity on User Search Behavior in Question Answering Systems', *LWA*, vol. 1/2006 of *Hildesheimer Informatik-Berichte*, pp. 132-137.
- [3] Antovski, L., and Armenski, G. (2012) 'KupiKniga.mk: transforming a website into a profitable e-commerce system using assisted conversions funnel', *ICT Innovations 2012 Web Proceedings*, ICT ACT, Skopje, pp. 321-329.
- [4] Arapakis, I., Lalmas, M., Cambazoglu, B. B., Marcos, M. C., and Jose, J. M. (2014) 'User engagement in online News: Under the scope of sentiment, interest, affect, and gaze', *Journal of the Association for Information Science and Technology*, 65(10), 1988-2005.
- [5] Attfield, S., Kazai, G., Lalmas, M., and Piwowarski, B. (2011, February) 'Towards a science of user engagement (position paper)', In *WSDM Workshop on User Modelling for Web Applications*.
- [6] Aula, A., Khan, R. M., and Guan, Z. (2010, April) 'How does search behavior change as search becomes more difficult?', In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 35-44, ACM.
- [7] Awad, M., and Khalil, I. (2012) 'Prediction of user's web-browsing behavior: Application of markov model', *Systems, Man, and Cybernetics, Part B: Cybernetics*, *IEEE Transactions on*, 42(4), 1131-1142.
- [8] Baeza-Yates, R., Calderón-Benavides, L., and González-Caro, C. (2006, January) 'The intention behind web queries', In *String processing and information retrieval*, pp. 98-109, Springer Berlin Heidelberg.
- [9] Baeza-Yates, R., Hurtado, C., Mendoza, M. and Dupret, G. (2005) 'Modelling User Search Behavior', *LA-WEB '05: Proceedings of the Third Latin American Web Congress*.
- [10] Bahadori, M., Harounabadi, A., Sadeghzadeh, M. (2013) 'Modeling user navigation behavior in web by colored Petri nets to determine the user's interest in recommending web pages', *Management Science Letters*, Volume 3 Issue 1, pp. 359-366.
- [11] Bause, F., and Kritzinger, P. S. (1996) 'Stochastic Petri Nets', *Vieweg+ Teubner Verlag*, pp: 163-171.

- [12] Birgit, H., Wets, G., and Vanhoof, K., (2001) 'Clustering navigation patterns on a website using a sequence alignment method', *Intelligent Techniques for Web Personalization: IJCAI* 1-6.
- [13] Bonet, P., Lladó, C. M., Puijaner, R., and Knottenbelt, W. J. (2007, October) 'PIPE v2. 5: A Petri net tool for performance modelling', In *Proc. 23rd Latin American Conference on Informatics (CLEI 2007)*.
- [14] Catledge, L.D., and Pitkow J., (1995) 'Characterizing browsing strategies in the World-Wide Web', *Computer Networks and ISDN systems* 27.6, pp. 1065-1073.
- [15] Chao, C., and Thomaz, A., (2012) 'Timed petri nets for multimodal interaction modeling', In *ICMI 2012 Workshop on Speech and Gesture Production in Virtually and Physically Embodied Conversational Agents*.
- [16] Chen Z, Lin F., Liu H., Liu Y., Wenyin L. and Ma W. (2002) 'User Intention Modeling in Web Applications Using Data Mining', *Kluwer Academic Publishers Hingham, Journal World Wide Web*, vol. 5, pp. 181-191, MA, USA.
- [17] Chiola, G., Franceschinis, G., Gaeta, R., and Ribauda, M. (1995), 'GreatSPN 1.7: graphical editor and analyzer for timed and stochastic Petri nets', *Performance evaluation*, 24(1), pp: 47-68.
- [18] Ciardo, G. and Lindemann, C. (1993) 'Analysis of deterministic and stochastic Petri nets', *Petri Nets and Performance Models*, 1993. *Proceedings*, 5th International Workshop on Digital Object Identifier, 10.1109/PNPM.1993.393454, pp. 160-169.
- [19] Ciardo, G., Muppala, J. and Trivedi, T. (1989) 'SPNP: stochastic Petri net package', *Petri Nets and Performance Models. PNPM89*, *Proceedings of the Third International Workshop on Digital Object Identifier*, 10.1109/PNPM.1989.68548, pp. 142-151.
- [20] Dingle, N. J., Knottenbelt, W. J., and Suto, T. (2009) 'PIPE2: a tool for the performance evaluation of generalised stochastic Petri Nets', *ACM SIGMETRICS Performance Evaluation Review*, 36(4), pp. 34-39.
- [21] Dobrian, F., Sekar, V., Awan, A., Stoica, I., Joseph, D., Ganjam, A., and Zhang, H. (2011) 'Understanding the impact of video quality on user engagement', *ACM SIGCOMM Computer Communication Review*, 41(4), pp. 362-373.
- [22] Dobrian, F., Awan, A., Joseph, D., Ganjam, A., Zhan, J., Sekar, V., and Zhang, H. (2013) 'Understanding the impact of video quality on user engagement'. *Communications of the ACM*, 56(3), pp: 91-99.
- [23] Downey, D. (2007) 'Models of Searching and Browsing Languages Studies and Applications', *Proceeding IJCAI*, pp. 2740-2747.
- [24] Gega, V. and Mitrevski, P. (2012) 'On the General Principles of Human-Computer Information Retrieval', *ICT Innovations 2012, Web Proceedings*, pp. 371-380.

- [25] Geisler, C. (2013) 'Designing for User Engagement on the Web: 10 Basic Principles. Routledge'.
- [26] Held, C., Kimmerle, J., and Cress, U. (2012) 'Learning by foraging: The impact of individual knowledge and social tags on web navigation processes', *Computers in Human Behavior*, 28(1), pp: 34-40.
- [27] Hsieh-Yee, I. (2001) 'Research on Web search behavior', *Library and Information Science Research*, 23(2), pp. 167-185.
- [28] Järvelin, K. and Ingwersen, P. (2004) 'Information seeking research needs extension towards tasks and technology', *Information Research*, vol. 10, no. 1.
- [29] Jenamani, M., Mohapatra, P. K., and Ghose, S. (2003) 'A stochastic model of e-customer behavior', *Electronic Commerce Research and Applications*, 2(1), pp. 81-94.
- [30] Jensen, K. (2013) 'Coloured Petri nets: basic concepts, analysis methods and practical use', (Vol. 1), Springer Science & Business Media.
- [31] Kantor, P. B. and Nordlie, R. (1999) 'Models of the Behavior of People Searching the Internet A Petri Net Approach', Larry Woods (Ed.), *Proceedings of the 62nd ASIS Annual Meeting*.
- [32] Kouzehgar, M., Badamchizadeh, M., and Khanmohammadi, S. (2013) 'Fuzzy Petri Nets for Human Behavior Verification and Validation', arXiv preprint arXiv: 1303.1247.
- [33] Lacerda, A., and Ziviani, N. (2013, February) 'Building user profiles to improve user experience in recommender systems', In *Proceedings of the sixth ACM international conference on Web search and data mining*, pp. 759-764, ACM.
- [34] Lee, U., Liu, Z., and Cho, J. (2005, May) 'Automatic identification of user goals in web search', In *Proceedings of the 14th international conference on World Wide Web*, pp. 391-400, ACM.
- [35] Lehmann, J., Lalmas, M., Yom-Tov, E., and Dupret, G. (2012) 'Models of user engagement', In *User Modeling, Adaptation, and Personalization*, pp. 164-175, Springer Berlin Heidelberg.
- [36] Leino, J. (2014) 'User Factors in Recommender Systems: Case Studies in e-Commerce', *News Recommending, and e-Learning*.
- [37] Levene, M. (2011) 'An introduction to search engines and web navigation', John Wiley and Sons.
- [38] Liu, Y., and Liu, Z. Y. (2012) 'Web Design and Analysis of Usability Issues and Recommendations', In *Advances in Computer Science and Information Engineering*, pp. 609-615, Springer Berlin Heidelberg.



- [39] Luo, Y. (2012) 'E-commerce personalized recommendation system design based on multi-Agent', In *Advances in Computer Science and Information Engineering*, pp. 159-164, Springer Berlin Heidelberg.
- [40] MacLennan, J., and Tang, Z. (2005) 'Data Mining with SQL Server 2005', Wiley.
- [41] Mangal, D., and Arya, K. V. (2014) 'An efficient approach for web path traversal pattern based on visitor preferences and navigation behavior', In *Industrial and Information Systems (ICIIS), 2014 9th International Conference* pp: 1-5, IEEE.
- [42] Mitrevski, P.J. and Hristoski, I.S., (2014) 'Behavioral-based performability modeling and evaluation of e-commerce systems', *Electronic Commerce Research and Applications* 13(5), pp. 320-340.
- [43] Mobasher, B., Jain, N., Han, E. H., and Srivastava, J. (1996) 'Web mining: Pattern discovery from world wide web transactions', Technical Report TR96-050, Department of Computer Science, University of Minnesota.
- [44] Molloy, M. K. (1982) 'Performance analysis using stochastic Petri nets', *Computers, IEEE Transactions on*, 100(9), pp. 913-917.
- [45] Murata, T. (1989) 'Petri Nets: Properties, Analysis and Applications', *Proceedings of the IEEE*, Vol.77, No.4, pp. 541-580.
- [46] O'Brien, H. L., and Toms, E. G. (2008) 'What is user engagement? A conceptual framework for defining user engagement with technology', *Journal of the American Society for Information Science and Technology*, 59(6), pp. 938-955.
- [47] O'Brien, H. L., and Toms, E. G. (2010) 'The development and evaluation of a survey to measure user engagement', *Journal of the American Society for Information Science and Technology*, 61(1), pp. 50-69.
- [48] O'Brien, H. L. (2010) 'The influence of hedonic and utilitarian motivations on user engagement: The case of online shopping experiences', *Interacting with Computers*, 22(5), pp. 344-352.
- [49] O'Brien, H. L. (2011) 'Exploring user engagement in online news interactions', *Proceedings of the American Society for Information Science and Technology*, 48(1), pp. 1-10.
- [50] Pass, G., Chowdhury, A. and Torgeson, C., (2006) 'A picture of search', *Proceeding InfoScale '06 Proceedings of the 1st international conference on Scalable information systems*, Article no.1.
- [51] Peterson, J. L. (1977) 'Petri nets', *ACM Computing Surveys (CSUR)*, 9(3), pp: 223-252.
- [52] Peterson, E. T., and Carrabis, J. (2008) 'Measuring the immeasurable: Visitor engagement', In *Web Analytics Demystified*, 14, 16.

- [53] Pitkow, J. E., Schütze H., Cass, T. A., Cooley, R., Turnbull, D., Edmonds, A., Adar, E. and Breuel, T. M. (2002) 'Personalized search', *Commun. ACM*, vol. 45, no. 9, pp. 50-55.
- [54] Piwowarski, B. and Zaragoza, H. (2007) 'Predictive User Click Models Based on click-through history', *Proceedings of the sixteenth ACM Conference on information and knowledge management*, pp. 175-182, ACM New York, NY, USA.
- [55] Po-Zung, C., Sun, C., and Shih-Yang, Y. (2008) 'Modeling and analysis the web structure using stochastic timed Petri nets' *Journal of Software* 3.8, pp.19-26.
- [56] Priyanka, M., Gulati, P., and Sharma, A. (2010) 'A novel approach for predicting user behavior for improving web performance', *International Journal on Computer Science and Engineering* 2.04.
- [57] Radinsky, K., Svore, K. M., Dumais, S. T., Teevan, J., Bocharov, A. and Horvit, E. (2012) 'Modeling and predicting behavioral dynamics on the web', *WWW*, pp. 599-608.
- [58] Reisig, W. (2012) 'Petri nets: an introduction', (Vol. 4). Springer Science and Business Media.
- [59] Reisig, W., and Rozenberg, G. (1998) 'Lectures on petri nets i: basic models: advances in petri nets', (Vol. 1491). Springer Science and Business Media.
- [60] Rieh, S. Y., and Hong, I. X., (2006) 'Analysis of multiple query reformulations on the web: The interactive information retrieval context', *Information Processing and Management* 42.3, pp. 751-768.
- [61] Riivo, K., and Potisepp, K. (2009) 'Click-stream Data Analysis of Web Traffic'.
- [62] Rose, D. E., and Levinson, D. (2004, May) 'Understanding user goals in web search', *In Proceedings of the 13th international conference on World Wide Web*, pp. 13-19, ACM.
- [63] Ruthven, I. (2011) 'Information Retrieval in Context', *Advanced Topics in Information Retrieval, The Information Retrieval Series*, vol. 33, pp. 187-207.
- [64] Sadikov, E., Madhavan, J., Wang, L., and Halevy, A. (2010, April) 'Clustering query refinements by user intent', *In Proceedings of the 19th international conference on World wide web*, pp. 841-850, ACM.
- [65] Sarwar, B. M., Karypis, G., Konstan, J., and Riedl, J. (2002, December) 'Recommender systems for large-scale e-commerce: Scalable neighborhood formation using clustering', *In Proceedings of the fifth international conference on computer and information technology (Vol. 1)*.
- [66] Schafer, J. B., Konstan, J. A., and Riedl, J. (2001) 'E-commerce recommendation applications', *In Applications of Data Mining to Electronic Commerce*, pp. 115-153, Springer US.

- [67] Schafer, J. B., Konstan, J., and Riedl, J. (1999, November) 'Recommender systems in e-commerce', In Proceedings of the 1st ACM conference on Electronic commerce, pp. 158-166, ACM.
- [68] Shih-Yang, Y., Po-Zung, C., and Sun, C. (2007) 'Using Petri Nets to enhance web usage mining.' Acta Polytechnica Hungarica 4.3.
- [69] Silverstein, C., Marais, H., Henzinger, M. and Moricz, M. (1999) 'Analysis of a very large web search engine query log', ACM New York, NY, USA, vol. 33, pp. 6-12.
- [70] Spink, A., Bateman, J., and Jansen, B. J., (1998) 'Searching heterogeneous collections on the web: behaviour of excite users', Information Research 4.2 pp. 4-2.
- [71] Su, Q., and Chen, L. (2014) 'A method for discovering clusters of e-commerce interest patterns using click-stream data', Electronic Commerce Research and Applications.
- [72] Tadeusz, M., Wojciechowski, M., and Zakrzewicz, M., (2000) 'Web users clustering', Proc. of the 15th International Symposium on Computer and Information Sciences.
- [73] Tan, G. W., and Wei, K. K. (2007) 'An empirical study of Web browsing behaviour: Towards an effective Website design', Electronic Commerce Research and Applications, 5(4), pp. 261-271.
- [74] Van Der Aalst, W. (2011) 'Process mining: discovery, conformance and enhancement of business processes', Springer Science and Business Media.
- [75] Wang, J. (2007) 'Petri nets for dynamic event-driven system modeling', In Handbook of Dynamic System Modeling, pp: 1-17.
- [76] Wei, K., Huang, J., and Fu, S. (2007, June) 'A survey of e-commerce recommender systems', In Service Systems and Service Management, 2007 International Conference on, pp. 1-5, IEEE.
- [77] Weinan, W., and Zaïane, O. R., (2002) 'Clustering web sessions by sequence alignment', Database and Expert Systems Applications, Proceedings 13th International Workshop on IEEE.
- [78] White, R. W. and Drucker, S. M. (2007) 'Investigating Behavioral Variability in Web Search', Proceedings of the 16th international conference on World Wide Web, pp. 21-30, ACM New York, NY, USA.
- [79] Yongli, R., Tomko, M., Ong K., and Sanderson M. (2014) 'How People Use the Web in Large Indoor Spaces', Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management. ACM.
- [80] Yordanova, K. (2011) 'Toward a unified human behaviour modelling approach', Technical Report CS-02-11, Institut für Informatik, Universität Rostock, Rostock, Germany.

- [81] Ypma, A. and Heskes, T. (2002) 'Automatic Categorization of Web Pages and User Clustering with Mixtures of Hidden Markov Models', WEBKDD, pp. 35-49.
- [82] Yu, Y., and Wu, X. (2010, August) 'Research on user experience design of B2C e-commerce based on personalized needs', In Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2010 2nd International Conference on, Vol. 2, pp. 100-103, IEEE.
- [83] Zhang, Z., and Qian, S. (2012) 'The Research of E-commerce Recommendation System Based on Collaborative Filtering Technology', In Advances in Computer Science and Information Engineering, pp. 507-512, Springer Berlin Heidelberg.

## Листа со табели

---

Табела 1 Места во GSPN мрежата .....	26
Табела 2 Премини во GSPN мрежата .....	26
Табела 3 Можно сценарио на проширениот модел .....	29
Табела 4 Forward incidence матрицата $I^+$ .....	31
Табела 5 Backward incidence матрицата $I^-$ .....	31
Табела 6 Комбинирана incidence матрицата $I$ .....	31
Табела 7 Интензитет на временските премини .....	33
Табела 8 Веројатности на активација.....	33
Табела 9 Просечно време на престој .....	34
Табела 10 Листа на страни по категории .....	36
Табела 11 Места во GSPN мрежата.....	38
Табела 12 Премини во GSPN мрежата.....	38
Табела 13 Означувања во GSPN мрежата .....	48
Табела 14 Forward incidence матрицата $I^+$ .....	49
Табела 15 Backward incidence матрицата $I^-$ .....	50
Табела 16 Комбинирана incidence матрицата $I$ .....	51
Табела 17 Преглед на податоците од статистичката анализа .....	54
Табела 18 Најчесто прегледувани страни (линкови) во лог датотеката.....	55
Табела 19 Најчесто посетувани страни (линкови) од уникатни корисници .....	56
Табела 20 Профили на кластери.....	60
Табела 21 Интензитет на временските премини пресметани за Кластер 1 .....	67
Табела 22 Веројатности на активација на премините во Кластер 1.....	68
Табела 23 Infinitesimal generator матрица за Кластер 1 .....	68
Табела 24 Просечно време на престој во состојбите за Кластер 1.....	69
Табела 25 Ограничување на Infinitesimal generator матрица за Кластер 1 .....	69
Табела 26 Очекувано време поминато во сите преодни допирливи означувања за Кластер 1 .....	69
Табела 27 Ограничување на веројатносната транзициска матрица за Кластер 1 .....	70
Табела 28 Просечен број на посети во сите преодни состојби во Кластер 1 .....	70
Табела 29 Кумулативното време на престој во сите преодни означувања во Кластер 1.....	70
Табела 30 Интензитет на временските премини пресметани за Кластер 2 .....	71
Табела 31 Веројатности на активација на премините во Кластер 1.....	71
Табела 32 Infinitesimal generator матрица за Кластер 2 .....	72
Табела 33 Просечно време на престој во состојбите за Кластер 2.....	72
Табела 34 Ограничување на Infinitesimal generator матрица за Кластер 1 .....	73
Табела 35 Очекувано време поминато во сите преодни допирливи означувања за Кластер 2 .....	73
Табела 36 Ограничување на транзициската веројатносна матрица за Кластер 2 .....	73
Табела 37 Просечен број на посети во сите преодни состојби во Кластер 2 .....	74

Табела 38 Кумулативното време на престој во сите преодни означувања во Кластер 2.....	74
Табела 39 Пример од лог податоците од процесот на пребарување.....	105
Табела 40 Краток пример од лог податоците од навигација.....	105
Табела 41 Популарност на категориите – целосна табела.....	107
Табела 42 Исечок од предвидувањето на кластерите .....	108
Табела 43 Преглед на податоците од статистичката анализа за Кластер 1 .....	109
Табела 44 Преглед на податоците од статистичката анализа за Кластер 2 .....	109
Табела 45 Транзициска веројатносна матрица за првиот кластер .....	110
Табела 46 Транзициска веројатносна матрица за вториот кластер.....	110

## Листа со слики

---

Слика 1 Пример за динамичкото однесување на едноставна Петриева мрежа, пред (лево) и после (десно) активирање на транзицијата $t$ .....	10
Слика 2 Пример за GSPN мрежа (лево) и соодветниот дијаграм на интензитетите на состојбите на премин на СТМС (десно) .....	13
Слика 3 Едноставен модел на корисничко однесување при пребарување .....	27
Слика 4 Проширен модел на корисничко однесување при пребарување .....	28
Слика 5 Граф на достигнување .....	30
Слика 6 Дијаграм на интензитетите на состојбите на премин на СТМС.....	31
Слика 7 Структура на веб страната.....	35
Слика 8 Модел на корисничко однесување при навигација.....	39
Слика 9 Граф на достигнување .....	48
Слика 10 Дијаграм на интензитетите на состојбите на премин на СТМС.....	52
Слика 11 Визуелен приказ на бројот на најчесто посетувани страни (линкови) .....	56
Слика 12 Визуелен приказ на бројот на најчесто посетувани линкови од уникатни корисници.....	57
Слика 13 Анализа на најпопуларни категории на книги (производи) .....	58
Слика 14 Број на предвидени корисници по кластер .....	60
Слика 15 Карактеристики на Кластер 1 .....	61
Слика 16 Карактеристики на Кластер 2 .....	62
Слика 17 Дел од разликите меѓу двата кластери .....	63
Слика 18 Дијаграм на преодни состојби во Кластер 1 .....	64
Слика 19 Дијаграм на преодни состојби во Кластер 2 .....	65
Слика 20 Најкарактеристични преодни состојби во Кластер 1.....	66
Слика 21 Најкарактеристични преодни состојби во Кластер 2.....	66
Слика 22 Просечен број на прегледи на страни по посета во Кластер 1 .....	79
Слика 23 Распределба на $C_i$ вредностите за 916 корисници и 7570 посети од Кластер 1 .....	79
Слика 24 Просеча должина на посетите во Кластер 1 .....	80
Слика 25 Распределба на $D_i$ вредностите за 916 корисници и 7570 посети од Кластер 1 .....	80
Слика 26 Просечена фреквенција на повторна посета во Кластер 1 .....	81
Слика 27 Распределба на $R_i$ вредностите за 916 корисници и 7570 посети од Кластер 1 .....	81
Слика 28 Просечен број посети во Кластер 1 .....	82
Слика 29 Распределба на $L_i$ вредностите за 916 корисници и 7570 посети од Кластер 1.....	82
Слика 30 Распределба на $I_i$ вредностите за 916 корисници и 7570 посети од Кластер 1 .....	83
Слика 31 Просечен број на прегледи на страни по посета во Кластер 2.....	83
Слика 32 Распределба на $C_i$ вредностите за 1068 корисници и 7863 посети од Кластер 2.....	84
Слика 33 Просеча должина на посетите во Кластер 2 .....	85
Слика 34 Распределба на $D_i$ вредностите за 1068 корисници и 7863 посети од Кластер 2.....	85
Слика 35 Просечена фреквенција на повторна посета во Кластер 2.....	85

Слика 36	Распределба на $R_i$ вредностите за 1068 корисници и 7863 посети од Кластер 2 .....	86
Слика 37	Просечен број посети во Кластер 2 .....	86
Слика 38	Распределба на $L_i$ вредностите за 1068 корисници и 7863 посети од Кластер 2 .....	87
Слика 39	Распределба на $I_i$ вредностите за 1068 корисници и 7863 посети од Кластер 2.....	87



## Додаток А – Ресурси

Во Табела 39 даден е пример како изгледа лог датотеката, односно поставените кориснички прашања при процесот на пребарување, собрани на клиентска страна. Додека пак во Табела 40 прикажан е исечок од лог датотеката генерирана на серверска страна, а се однесува на податоците собрани при процесот на навигација на корисниците.

**Табела 39 Пример од лог податоците од процесот на пребарување**

Searched for <a href="#">Computation of Steady-State Probability Distributions of CTMC</a>	2:58pm
<a href="#">Continuous-time Markov chains</a> <a href="#">cgi.di.uoa.gr</a>	
Searched for <a href="#">Computation of Steady-State Probability Distributions</a>	2:13pm
Searched for <a href="#">steady state probability distribution</a>	2:13pm
<a href="#">http://www.math.ncsu.edu/ma114/PDF/6.3.pdf</a>	
Searched for <a href="#">solve system of linear equations for ergodic CTMC</a>	2:08pm
<a href="#">A general algorithm to compute the steady-state solution of product ...</a> <a href="#">dsi.unive.it</a>	
<a href="#">t - IRCCyN</a> <a href="#">irccyn.ec-nantes.fr</a>	
Searched for <a href="#">memoryless property of continuous time markov chain</a>	12:44pm
<a href="#">Continuous Time Markov Chains (CTMCs)</a> <a href="#">win.tue.nl</a>	

**Табела 40 Краток пример од лог податоците од навигација**

138	31.11.99.152	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://kupikniga.mk/PrizeContest.aspx?Code=2</a>	24.03.2013 22:09:56
138	31.11.99.152	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://kupikniga.mk/PrizeContest.aspx?Code=2</a>	24.03.2013 22:10:00
138	31.11.99.152	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://kupikniga.mk/BookContest.aspx</a>	24.03.2013 22:10:00
138	31.11.99.152	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://www.kupikniga.mk/DeliveryPolitics.aspx</a>	24.03.2013 23:06:26
138	31.11.99.152	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://www.kupikniga.mk/BackPolicy.aspx</a>	24.03.2013 23:06:32
138	31.11.99.152	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://www.kupikniga.mk/HelpOrders.aspx</a>	24.03.2013 23:06:43
138	31.11.99.152	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://www.kupikniga.mk/Default.aspx</a>	24.03.2013 23:06:21
138	31.11.99.152	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://kupikniga.mk/BookDetails.aspx?Pr=31594</a>	28.03.2013 23:28:56
141	194.149.136.110	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://www.kupikniga.mk/MyLoungeDef.aspx</a>	19.03.2013 12:48:00
141	194.149.136.110	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://www.kupikniga.mk/Default.aspx</a>	19.03.2013 12:48:30
141	194.149.136.110	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://www.kupikniga.mk/Default.aspx</a>	19.03.2013 12:48:33
141	194.149.136.110	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://www.kupikniga.mk/SearchPreview.aspx</a>	19.03.2013 12:48:37
141	194.149.136.110	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://www.kupikniga.mk/SearchPreview.aspx</a>	19.03.2013 12:49:06
141	194.149.136.110	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://www.kupikniga.mk/SearchPreview.aspx</a>	19.03.2013 12:49:11
141	194.149.136.110	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://www.kupikniga.mk/SearchPreview.aspx</a>	19.03.2013 12:49:30
150	217.16.79.83	mk-MK	IE	IE9	9	<a href="#">http://kupikniga.mk/MyLoungeDef.aspx</a>	11.04.2013 14:01:58
159	82.214.112.20	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://kupikniga.mk/MyLoungeDef.aspx</a>	01.05.2013 12:09:16
159	82.214.112.20	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://kupikniga.mk/Default.aspx</a>	01.05.2013 12:09:17
159	82.214.112.20	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://kupikniga.mk/NewProducts.aspx</a>	01.05.2013 12:09:19
159	82.214.112.20	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://kupikniga.mk/BestSellers.aspx</a>	01.05.2013 12:09:20
159	88.85.115.11	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://kupikniga.mk/MyLoungeDef.aspx</a>	06.03.2013 12:07:18
159	88.85.115.11	en-US,en;q=0.8	AppleMAC-Safari	Desktop	5	<a href="#">http://kupikniga.mk/MyLoungeDef.aspx</a>	06.03.2013 12:06:52

## Додаток Б – Работна околина

---

За кодирање на скриптата со која беа парсирани, обработени и исчистени изворните лог датотеки во облик едноставен за понатамошна обработка и статистичка анализа, се користеше Microsoft Visual Studio 2013<sup>11</sup> и програмски јазик C#.

За дел од статистичките анализи кои не беа реализирани со скриптата, како и креирањето на графичите, се користеше Microsoft Excel 2010<sup>12</sup>.

Исто така, посложените математички равенки, беа решени со помош на Matlab<sup>13</sup>.

За цртање и анализа на Петриевите мрежи беше употребен Pipe Petri Net Editor<sup>14</sup> [13] и [20].

За реализација на делот што се однесува на кластрирањето секвенци, беше искористен SQL Server 2014 Bussines Inteligenca<sup>15</sup> Module и Bussiness Development Studio<sup>16</sup> [40].

---

<sup>11</sup> <https://msdn.microsoft.com/en-us/library/dd831853.aspx>

<sup>12</sup> <https://www.microsoft.com/en-us/download/details.aspx?id=16642>

<sup>13</sup> <http://www.mathworks.com/products/matlab/>

<sup>14</sup> <http://pipe2.sourceforge.net/>

<sup>15</sup> <http://www.microsoft.com/en-us/server-cloud/products/sql-server-editions/sql-server-business-intelligence.aspx>

<sup>16</sup> [https://msdn.microsoft.com/en-us/library/ms173767\(v=sql.105\).aspx](https://msdn.microsoft.com/en-us/library/ms173767(v=sql.105).aspx)

## Додаток Ц – Статистички анализи и Евалуација на перформанси

Во Табела 41 прикажани се сите категории на книги присутни во периодот во кој се регистрирани лог податоците и нивната популарност.

Табела 41 Популарност на категориите – целосна табела

Fiction	24.75%	Драма (трагедии)	0.43%
Popular Psychology	7.55%	Земјоделство	0.38%
Novels	6.49%	Општествени науки	0.38%
Esoteric	4.63%	Поезија	0.30%
Children's books	4.24%	Туризам	0.28%
Philosophy and Religion	3.61%	Сложувалки и игри	0.27%
Love	3.00%	Готвачи	0.27%
Dictionaries	2.40%	Сликовници	0.26%
History	2.36%	Економија	0.26%
Triller	2.26%	Машинство	0.25%
Business	2.13%	Популарна наука	0.24%
Atlases and encyclopedias	2.12%	Карти и мапи	0.24%
Professional and Technical	2.12%	Сребрен накит	0.24%
Other	2.06%	Електротехника	0.22%
Informatics	1.89%	Музика	0.22%
Scifi	1.65%	Детски енциклопедии	0.22%
Health	1.45%	Маслиново дрво	0.20%
Sets	1.39%	ДВД	0.20%
Drama	1.33%	Авантура	0.19%
Life	1.27%	Математика	0.18%
Biography	1.16%	Уметност	0.16%
Teen	1.13%	Географија, историја и археологија	0.14%
Comics	1.06%	Учебници	0.13%
Adults books 18+	1.04%	Руски класици	0.12%
Classical literature	1.02%	Социологија и антропологија	0.12%
Literature Theory	0.92%	Раскази	0.12%
Medicine	0.86%	Природни науки	0.11%
School reading	0.80%	Лектири за средно образование	0.11%
Law	0.76%	Есеи и колумни	0.10%
Jewelry	0.74%	Книги со стикери	0.10%
Mistery	0.70%	Бронзени предмети	0.10%
Criminal	0.60%	Мултимедија	0.10%
Souvenirs	0.53%	Од нашата традиција	0.09%
Монографии	0.48%	Научна Психологија и Психијатрија	0.08%
Политика	0.46%	Спорт	0.07%
Архитектура и градежништво	0.45%	Екологија	0.07%
Психолошки	0.44%		

Историја	0.07%	Петто одделение	0.02%
Политикологија	0.07%	Шесто одделение	0.02%
Бронзен накит	0.06%	Цртани филмови - DVD	0.02%
Маркетинг	0.06%	Фармација	0.02%
Менаџмент	0.05%	Логистика	0.02%
Боенки	0.05%	Криминологија	0.02%
Второ одделение	0.05%	Претприемништво	0.02%
Осмо одделение	0.05%	Геологија	0.02%
Телекомуникации	0.05%	Статистика	0.01%
Сребрен накит за миленици	0.04%	Астрологија и хороскоп	0.01%
Биотехнологија	0.04%	Образование	0.01%
Графички дизајн	0.04%	Тарот карти	0.01%
Ботаника	0.04%	Астрономија	0.01%
Хорор	0.04%	Биологија	0.01%
Мода и дизајн	0.04%	Етнологија и етнографија	0.01%
Трето одделение	0.03%	Хемија	0.00%
Финансии и Сметководство	0.03%	Театар	0.00%
Прво одделение	0.03%	Етномузикологија	0.00%
Физика	0.03%	Текстил	0.00%
Седмо одделение	0.03%	Зоологија	0.00%
Спорт и тренинг	0.03%	Историски списанија	0.00%
Возачки испит	0.03%	Царина-шпедиција	0.00%
Четврто одделение	0.03%	На албански/Gjuha shqipe	0.00%
Ветерина	0.03%	Акција	0.00%

Во Табела 42 прикажан е сегмент од табелата во која ги содржи податоците од кластерирањето на корисниците во еден од предвидените кластери.

**Табела 42 Исечок од предвидувањето на кластерите**

Корисник	Предвиден кластер
138	Cluster 1
141	Cluster 2
150	Cluster 2
159	Cluster 1
161	Cluster 2
163	Cluster 2
170	Cluster 1
173	Cluster 2
175	Cluster 1
177	Cluster 1
178	Cluster 2
193	Cluster 1
208	Cluster 2

213	Cluster 2
214	Cluster 1
220	Cluster 2
227	Cluster 2
229	Cluster 1
230	Cluster 2
235	Cluster 1
236	Cluster 2
239	Cluster 1
245	Cluster 2
254	Cluster 1
294	Cluster 1
296	Cluster 1
304	Cluster 2
308	Cluster 1

309	Cluster 1
317	Cluster 1
329	Cluster 1
330	Cluster 1
333	Cluster 1
338	Cluster 1
362	Cluster 1
363	Cluster 1
368	Cluster 1
369	Cluster 1

375	Cluster 1
389	Cluster 1
392	Cluster 1
400	Cluster 1
406	Cluster 2
416	Cluster 1
432	Cluster 2
439	Cluster 2
442	Cluster 1

Податоците кои се однесуваат на: времетраењето на корисничката сесија, бројот на сесии по корисник и бројот на посети по сесија, за секој кластер посебно, илустрирани се во Табела 43 и Табела 44.

Табела 43 Преглед на податоците од статистичката анализа за Кластер 1

Статистичка анализа на лог податоците				
Вкупен број на прегледи	98884			
Вкупен број на корисници	916			
Вкупен број на посети	7570			
	<b>Mean</b>	<b>(Min, Max)</b>		<b>SD</b>
Време помеѓу две последователни посети	5d:01h:11m:03s	(0d:0h:0m:0s, 301d:02h:42m:40s)		
Број на посети по корисник	8.26	(1, 58)		8.37
Број на прегледи по посета	13.06	(1, 441)		24.42

Табела 44 Преглед на податоците од статистичката анализа за Кластер 2

Статистичка анализа на лог податоците				
Вкупен број на прегледи	229970			
Вкупен број на корисници	1068			
Вкупен број на посети	7863			
	<b>Mean</b>	<b>(Min, Max)</b>		<b>SD</b>
Време помеѓу две последователни посети	2d:16h:07m:55s	(0d:0h:0m:0s, 332d:23h:52m:57s)		
Број на посети по корисник	7.36	(1, 229)		17.27
Број на прегледи по посета	28.36	(1, 2929)		68.86

Во Табела 45 и Табела 46, прикажани се транзициските веројатносни матрици од двата кластери, кои учествуваат во пресметките во секција 4.2.4.3.1 и 4.2.4.3.2, соодветно.

**Табела 45 Транзициска веројатносна матрица за првиот кластер**

	M <sub>A</sub>	M <sub>B</sub>	M <sub>E</sub>	M <sub>L</sub>	M <sub>ML</sub>	M <sub>C</sub>	M <sub>D1</sub>	M <sub>D2</sub>	M <sub>D3</sub>	M <sub>D4</sub>
M <sub>A</sub>	0.000000	0.426888	0.021765	0.025955	0.000000	0.000000	0.525392	0.000000	0.000000	0.000000
M <sub>B</sub>	0.564919	0.000000	0.198443	0.236638	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
M <sub>E</sub>	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
M <sub>L</sub>	0.000000	0.000000	0.027470	0.000000	0.972530	0.000000	0.000000	0.000000	0.000000	0.000000
M <sub>ML</sub>	0.000000	0.000000	0.003408	0.000000	0.000000	0.996592	0.000000	0.000000	0.000000	0.000000
M <sub>C</sub>	0.048362	0.333201	0.016989	0.000000	0.601448	0.000000	0.000000	0.000000	0.000000	0.000000
M <sub>D1</sub>	0.000103	0.000708	0.000036	0.000043	0.001278	0.000000	0.000000	0.997832	0.000000	0.000000
M <sub>D2</sub>	0.000180	0.001237	0.000063	0.000000	0.002234	0.000000	0.000000	0.000000	0.996286	0.000000
M <sub>D3</sub>	0.000145	0.001002	0.000051	0.000000	0.001809	0.000000	0.000000	0.000000	0.000000	0.996992
M <sub>D4</sub>	0.048362	0.333201	0.016989	0.000000	0.601448	0.000000	0.000000	0.000000	0.000000	0.000000

**Табела 46 Транзициска веројатносна матрица за вториот кластер**

	M <sub>A</sub>	M <sub>B</sub>	M <sub>E</sub>	M <sub>L</sub>	M <sub>ML</sub>	M <sub>C</sub>	M <sub>D1</sub>	M <sub>D2</sub>	M <sub>D3</sub>	M <sub>D4</sub>
M <sub>A</sub>	0.000000	0.975715	0.000341	0.000379	0.000000	0.000000	0.023565	0.000000	0.000000	0.000000
M <sub>B</sub>	0.840319	0.000000	0.075710	0.083971	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
M <sub>E</sub>	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
M <sub>L</sub>	0.000000	0.000000	0.036333	0.000000	0.963667	0.000000	0.000000	0.000000	0.000000	0.000000
M <sub>ML</sub>	0.000000	0.000000	0.002928	0.000000	0.000000	0.997072	0.000000	0.000000	0.000000	0.000000
M <sub>C</sub>	0.003831	0.986670	0.000345	0.000000	0.009154	0.000000	0.000000	0.000000	0.000000	0.000000
M <sub>D1</sub>	0.000570	0.146759	0.000051	0.000057	0.001362	0.000000	0.000000	0.851201	0.000000	0.000000
M <sub>D2</sub>	0.000521	0.134085	0.000047	0.000000	0.001244	0.000000	0.000000	0.000000	0.864103	0.000000
M <sub>D3</sub>	0.001777	0.457618	0.000160	0.000000	0.004246	0.000000	0.000000	0.000000	0.000000	0.536199
M <sub>D4</sub>	0.003831	0.986670	0.000345	0.000000	0.009154	0.000000	0.000000	0.000000	0.000000	0.000000

## Кратенки и акроними

---

Во овој додаток презентирани се дел од стручната терминологија која се користи во оваа докторска дисертација.

AOL – America Online

Bi – Brand Index

Ci – ClickDepth Index

CPN – Coloured Petri Net

CTMC – Continious Time Markov Chain

Di – Duration Index

Fi – Feedback Index

GSPN – Generalized Stochastic Petri Net

HMM – Hidden Markov Models

Ii – Interaction Index

Li – Loyalty Index

MC – Markov Chain

PN – Petri Net

Ri – Recency Index

SAM – Sequence Alignment Method

SPN – Stochastic Petri Net

STPN – Stochastic Timed Petri Net

W3C - World Wide Web Consortium